```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         May 12, 2003, 10:35:25 ; Search time 5814 Seconds
                                          (without alignments)
                                          7273.199 Million cell updates/sec

Title:          US-09-759-130B-381
Perfect score:  8138
Sequence:       1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:     4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09759130/runat_06052003_123554_23353/app_query.fasta_1.1607
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759130_@CGN_1_1_3897_@runat_06052003_123554_23353 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	8109	99.6	4579	9	AF264014	AF264014 Homo sapi
	2	6107.5	75.0	3670	6	AX133829	AX133829 Sequence
	3	3779	46.4	2249	9	AK058038	AK058038 Homo sapi
	4	3441.5	42.3	3786	9	HSM130AC1	Z22969 H.sapiens m
	5	3429	42.1	3703	9	HSM130A	Z22968 H.sapiens m
	6	3429	42.1	4950	9	HSM130AC2	Z22970 H.sapiens m
	7	3396.5	41.7	3802	9	HSM130AE	Z22971 H.sapiens m
	8	3377	41.5	3707	6	AX281768	AX281768 Sequence
	9	3272	40.2	4405	10	AF274883	AF274883 Mus muscu
	10	3082	37.9	4475	4	BBWC11MR	X63723 B.bovis WC1
	11	2990	36.7	7656	9	HSA243224	AJ243224 Homo sapi
	12	2987	36.7	7683	9	HSA243212	AJ243212 Homo sapi
	13	2982	36.6	7667	9	HSA297935	AJ297935 Homo sapi
	14	2980	36.6	7686	9	AF159456	AF159456 Homo sapi
c	15	2783	34.2	202252	2	AC131207	AC131207 Homo sapi
	16	2447.5	30.1	7006	3	AF064259	AF064259 Strongylo
	17	2435	29.9	5802	6	A90822	A90822 Sequence 1
	18	2435	29.9	5802	6	AR193170	AR193170 Sequence
	19	2435	29.9	5802	9	HSDMBT1	AJ000342 Homo sapi
	20	2246.5	27.6	6323	3	GCY14953	Y14953 Geodia cydo
	21	2161	26.6	5108	4	AF043112	AF043112 Oryctolag
	22	1963.5	24.1	6652	10	MMU37438	U37438 Mus musculu
	23	1856.5	22.8	6290	10	AB005909	AB005909 Mus muscu
	24	1594.5	19.6	3769	3	AF228827	AF228827 Strongylo
	25	1586	19.5	2305	4	AF400115	AF400115 Ovis arie
c	26	1465	18.0	170882	2	AC107839	AC107839 Mus muscu
	27	1442	17.7	3431	3	AF228826	AF228826 Strongylo
	28	1399	17.2	2423	4	SSSRP5	X99336 S.scrofa mR
	29	1341	16.5	1665	4	SSSRP3	X99334 S.scrofa mR
	30	1305.5	16.0	177768	2	AC131206	AC131206 Homo sapi
	31	1304	16.0	3927	3	AF228824	AF228824 Strongylo
	32	1302	16.0	205331	2	AC124470	AC124470 Mus muscu
	33	1277	15.7	3332	3	AF228825	AF228825 Strongylo
	34	1263	15.5	170882	2	AC107839	AC107839 Mus muscu
	35	1205	14.8	2165	4	AF400122	AF400122 Ovis arie
	36	1204	14.8	1316	4	SSSRP2	X99333 S.scrofa mR
	37	1176	14.5	161644	9	AL161645	AL161645 Human DNA
	38	1142.5	14.0	4344	6	AX402036	AX402036 Sequence
	39	1142.5	14.0	4344	10	RNU32681	U32681 Rattus norv
	40	1142.5	14.0	4360	6	I72544	I72544 Sequence 1
	41	1085.5	13.3	1519	4	SSSRP4	X99335 S.scrofa mR
	42	1078.5	13.3	1800	4	S78981	S78981 mucin {clon
	43	1041.5	12.8	1654	5	AB051832	AB051832 Gallus ga
	44	968.5	11.9	2806	9	HSA312755	AJ312755 Homo sapi
c	45	916	11.3	192430	2	AC079542	AC079542 Mus muscu

ALIGNMENTS

RESULT 1

```
AF264014
LOCUS       AF264014              4579 bp    mRNA    linear   PRI 28-NOV-2000
DEFINITION  Homo sapiens scavenger receptor cysteine-rich type 1 protein M160
            precursor, mRNA, complete cds, alternatively spliced.
ACCESSION   AF264014
VERSION     AF264014.1  GI:9652086
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4579)
  AUTHORS   Gronlund,J., Vitved,L., Lausen,M., Skjodt,K. and Holmskov,U.
  TITLE     Cloning of a novel scavenger receptor cysteine-rich type I
            transmembrane molecule (M160) expressed by human macrophages
  JOURNAL   J. Immunol. 165 (11), 6406-6415 (2000)
  MEDLINE   20540099
   PUBMED   11086079
REFERENCE   2  (bases 1 to 4579)
  AUTHORS   Groenlund,J., Vitved,L., Lausen,M., Skjodt,K. and Holmskov,U.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2000) Department of Immunology & Microbiology,
            Institute of Medical Biology, University of Southern Denmark,
            Odense University, Winsloewparken 21.1, Odense 5000, Denmark
FEATURES             Location/Qualifiers
     source          1..4579
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q13.3"
     5'UTR           1..25
     CDS             26..4387
                     /codon_start=1
                     /product="scavenger receptor cysteine-rich type 1 protein
                     M160 precursor"
                     /protein_id="AAF91396.1"
                     /db_xref="GI:9652087"
                     /translation="MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGT
                     DLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQ
                     AVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG
                     NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIW
                     LDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVE
                     LKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGN
                     ESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWT
                     ICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTY
                     DGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTR
                     NAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCV
                     HREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQL
                     DCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVIC
                     SDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVS
                     REPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLV
                     GADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGL
                     TWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVE
                     INVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN
                     ESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALIC
                     LEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVS
                     AHFGEGSGPIWLDDLNCTGMESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSE
                     TETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGF
                     MWVDDIQCPKTHISIWQCLSAPWERRISSTAEETWITCEDRIRVRGGDTECSGRVEIW
                     HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES
                     FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGRLALILSSIFGLLLPVLFIL
                     FLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCED
                     ASDTSLLGVLPASEATK"
     sig_peptide     26..151
     misc_feature    152..469
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 1"
     misc_feature    470..790
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 2"
     misc_feature    791..1111
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 3"
     misc_feature    1112..1432
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 4"
     misc_feature    1433..1753
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 5"
     misc_feature    1754..2074
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 6"
     misc_feature    2075..2395
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 7"
     misc_feature    2396..2710
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 8"
     misc_feature    2711..3025
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 9"
     misc_feature    3026..3118
                     /note="Region: interdomain segment"
     misc_feature    3119..3433
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 10"
     misc_feature    3434..3754
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 11"
     misc_feature    3755..4063
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 12"
     misc_feature    4064..4093
                     /note="Region: connecting peptide"
     misc_feature    4094..4171
                     /note="Region: transmembrane domain"
     misc_feature    4172..4384
                     /note="Region: cytoplasmatic domain"
     3'UTR           4388..4579
     polyA_signal    4456..4461
     polyA_signal    4550..4555
BASE COUNT     1147 a    919 c   1343 g   1170 t
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:         4579
Score:                  8109.00        Matches:        1449
Percent Similarity:     99.72%         Conservative:   0
Best Local Similarity:  99.72%         Mismatches:     4
Query Match:            99.64%         Indels:         0
DB:                     9              Gaps:           0

US-09-759-130B-381 (1-1453) x AF264014 (1-4579)

Qy        1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       26 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 85

Qy       21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       86 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 145

Qy       41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      146 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 205

Qy       61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      206 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 265

Qy       81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      266 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCCTTCGCCATGTTTCGTTTT 325

Qy      101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    326 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 385

Qy    121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 445

Qy    141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    446 GAAGATGTTGGTGTGAACTGTTACGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 505

Qy    161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGAACTATATGT 565

Qy    181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    566 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 625

Qy    201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    626 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 685

Qy    221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 745

Qy    241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    746 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 805

Qy    261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    806 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 865

Qy    281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    866 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 925

Qy    301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    926 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 985

Qy    321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    986 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1045

Qy    341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db.  1046 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1105

Qy    361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1106 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1165

Qy    381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1166 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1225

Qy    401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1226 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1285

Qy    421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1286 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1345

Qy    441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1346 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1405

Qy    461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1406 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1465

Qy    481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1466 GCTCATAGCCCCTGTTATGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1525

Qy    501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1526 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1585

Qy    521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1586 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1645

Qy    541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1646 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1705

Qy    561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1706 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1765

Qy    581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1766 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1825

Qy    601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1826 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1885

Qy    621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1886 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1945

Qy    641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1946 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2005

Qy    661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2006 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2065

Qy    681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2066 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2125

Qy    701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2126 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2185

Qy    721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2186 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2245

Qy    741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2305

Qy    761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2365

Qy    781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2366 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2425

Qy    801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2485

Qy    821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTGTGGAGAT 2545
```

```
Qy    841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2546 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2605

Qy    861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2665

Qy    881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2666 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2725

Qy    901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2726 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2785

Qy    921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2786 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2845

Qy    941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2905

Qy    961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2965

Qy    981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3025

Qy   1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3085

Qy   1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3145

Qy   1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3205

Qy   1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3206 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3265

Qy   1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3325

Qy   1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
          |||||||||||||||||||||||     |||||||||||||||||||||||||||||||||
Db   3326 GATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3385

Qy   1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3386 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3445

Qy   1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3446 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3505

Qy   1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3506 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3565

Qy   1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3566 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3625

Qy   1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3626 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3685

Qy   1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
          ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db   3686 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCACAGCAGAAGAGACCTGGATC 3745

Qy   1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3746 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3805

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3806 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3865

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3866 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3925

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3926 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3985

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3986 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4045

Qy   1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db   4046 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCGTTTA 4105

Qy   1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
          ||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db   4106 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCCGGTTCTGTTTATTCTATTTCTC 4165

Qy   1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4166 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4225

Qy   1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4226 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4285

Qy   1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4286 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4345

Qy   1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
          |||||||||||||||||||||||||||||||||||||||
Db   4346 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4384
```

```
RESULT 2
AX133829
LOCUS       AX133829                 3670 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 15 from Patent WO0119856.
ACCESSION   AX133829
VERSION     AX133829.1  GI:14139781
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3670)
  AUTHORS   Shimkets,R.A., Fernandes,E., Herrmann,J.L., Liu,X., Yang,M. and
            Boldog,F.L.
  TITLE     Secreted human proteins, polynucleotides encoding them and methods
            of using the same
  JOURNAL   Patent: WO 0119856-A 15 22-MAR-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..3670
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
     CDS             113..3475
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC38965.1"
                     /db_xref="GI:14139782"
                     /translation="MKNDADVKIRLENRPNSENRCVWPTHRKMDGADLELRLADGSNN
                     CSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINS
                     ISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKY
                     QGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESN
                     IWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCD
                     DGWNGKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGW
                     GNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAE
                     VVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNM
                     EASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDA
                     ISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTD
                     VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGE
                     RSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVS
                     DPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVV
                     CQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGMESHLWQCPSRGWGQHDCRHKEDA
                     GVICSEFTALRLYSETETGSCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGEN
                     GVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIR
                     VRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGT
                     GTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALI
                     LSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPH
                     GTRTSDDTPNHGCEDASDTSLLGVLPASEATK"
BASE COUNT      947 a    752 c   1073 g    898 t
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          3670
Score:                    6107.50        Matches:         1103
Percent Similarity:       95.33%         Conservative:    0
Best Local Similarity:    95.33%         Mismatches:      3
Query Match:              75.05%         Indels:          51
DB:                       6              Gaps:            1

US-09-759-130B-381 (1-1453) x AX133829 (1-3670)

Qy    348 AsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer--------------- 362
          |||||||||||||||||||||||||||||||||||||||||||||
Db      2 AATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGACCAGTTGCTATTC 61

Qy    362 ------------------------------------------------------------ 362

Db     62 AAACTCTGGAACTTCTGGAACTATTTGCCTCTGGAAGAGTAAGGAAGGGCCATGAAGAAT 121

Qy    362 ------------------------------------------------------------ 362

Db    122 GATGCAGATGTGAAAATAAGGTTGGAAAACCGTCCCAACAGTGAAAACAGATGTGTATGG 181

Qy    363 ------------------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsn 376
                            ||||||||||||||||||||||||||||||||||||||||||
Db    182 CCAACCCACAGGAAAATGGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAAC 241

Qy    377 AsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGln 396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 AATTGTTCAGGAAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAG 301

Qy    397 AsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerVal 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 AACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTC 361

Qy    417 PheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIle 436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 TTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATA 421

Qy    437 SerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArg 456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 TCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGA 481

Qy    457 ThrCysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeu 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 ACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTA 541

Qy    477 ArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGlu 496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 AGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAG 601

Qy    497 TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln 516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 TGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAA 661

Qy    517 LeuGlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGly 536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 TTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGA 721

Qy    537 ProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu 556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722 CCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAA 781

Qy    557 HisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer 576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    782 CACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCA 841

Qy    577 GlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeu 596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    842 GGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTG 901

Qy    597 GluValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAla 616
          |||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db    902 GAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACGGTAAAGCT 961

Qy    617 AlaAlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGly 636
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    962 GCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGA 1021

Qy    637 AsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGlu 656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1022 AACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAG 1081

Qy    657 SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu 676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1082 TCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAA 1141

Qy    677 AspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySer 696
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1142 GATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGC 1201

Qy    697 SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla 716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1202 AGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCT 1261

Qy    717 AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla 736
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1262 AATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCA 1321

Qy    737 IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn 756
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1322 ATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAAT 1381

Qy    757 SerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln 776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1382 TCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAG 1441

Qy    777 ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArg 796
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1442 ACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGG 1501

Qy    797 LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp 816
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1502 CTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGG 1561

Qy    817 ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu 836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1562 CGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTA 1621
```

```
Qy      837 AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu 856
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1622 AACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTA 1681

Qy      857 ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle 876
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1682 ACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATT 1741

Qy      877 ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg 896
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1742 GTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGA 1801

Qy      897 TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1802 TATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAAC 1861

Qy      917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1862 GTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTT 1921

Qy      937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1922 CTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGA 1981

Qy      957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1982 GAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTG 2041

Qy      977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2042 GATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCT 2101

Qy      997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2102 GTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGAC 2161

Qy     1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2162 CCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTC 2221

Qy     1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2222 CGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTC 2281

Qy     1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2282 TGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAG 2341

Qy     1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2342 CTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGG 2401

Qy     1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
            ||||||||||||||||||||||||||||||||||||     |||||||||||||||||||
Db     2402 CCCATCTGGCTGGATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCT 2461

Qy     1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2462 TCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCA 2521

Qy     1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
            |||||||||||||||||||||||||||||||||||||||     ||||||||||||||||
Db     2522 GAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGGGAGCTGTGCTGGGAGATTG 2581

Qy     1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2582 GAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATA 2641

Qy     1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2642 GCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCT 2701

Qy     1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2702 TTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCAT 2761

Qy     1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2762 ATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAA 2821

Qy     1237 GluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSer 1256
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2822 GAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCT 2881

Qy     1257 GlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAsp 1276
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2882 GGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGAC 2941

Qy     1277 LeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeu 1296
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2942 CTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTG 3001

Qy     1297 ArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLys 1316
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3002 AGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAA 3061

Qy     1317 GlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGly 1336
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3062 GGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGA 3121

Qy     1337 HisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSer 1356
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3122 CACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCC 3181

Qy     1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3182 TCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTT 3241

Qy     1377 IleLeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSer 1396
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3242 ATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCA 3301

Qy     1397 ThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLys 1416
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3302 ACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAG 3361

Qy     1417 ArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAsp 1436
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3362 AGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGAT 3421

Qy     1437 AlaSerAspThrSerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3422 GCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 3472

RESULT 3
AK058038
LOCUS       AK058038                2249 bp    mRNA    linear   PRI 31-OCT-2001
DEFINITION  Homo sapiens cDNA FLJ25309 fis, clone SYN00985, highly similar to
            Homo sapiens scavenger receptor cysteine-rich type 1 protein M160
            precursor.
ACCESSION   AK058038
VERSION     AK058038.1  GI:16554040
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens synovial membrane (knee) cDNA to mRNA, clone_lib:SYN
            clone:SYN00985.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
            Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
            Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
            Yamashita,H., Chiba,Y., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S.,
            Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
            Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
            Sugano,S.
```

```
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2249)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
            full insert sequencing: RAB and Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .2249
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="SYN00985"
                     /tissue_type="synovial membrane (knee)"
                     /clone_lib="SYN"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      559 a    512 c    638 g    540 t
ORIGIN

Alignment Scores:
Pred. No.:              1.38e-237      Length:          2249
Score:                  3779.00        Matches:         682
Percent Similarity:     99.71%         Conservative:    0
Best Local Similarity:  99.71%         Mismatches:      2
Query Match:            46.44%         Indels:          0
DB:                     9              Gaps:            0

US-09-759-130B-381 (1-1453) x AK058038 (1-2249)

Qy      770 IleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCys 789
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3 ATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGC 62

Qy      790 SerAlaHisArgGlnProArgLeuValGlyAlaAspMetProCysSerGlyArgValGlu 809
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63 TCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAA 122

Qy      810 ValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAla 829
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      123 GTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCC 182

Qy      830 AsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHis 849
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      183 AATGTGCTGTGCAGAGAATTAAACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCAC 242

Qy      850 PheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGluGlySerGluThr 869
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      243 TTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACT 302

Qy      870 HisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIleHisSerArgGlu 889
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      303 CACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAA 362

Qy      890 ValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGlyLysSerGlnCys 909
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      363 GTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGT 422

Qy      910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      423 GACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGG 482

Qy      930 AspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      483 GACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACC 542

Qy      950 ThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCys 969
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      543 ACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGC 602

Qy      970 LeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCys 989
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      603 TTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGT 662

Qy      990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      663 ATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCA 722

Qy     1010 CysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIle 1029
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      723 TGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATC 782

Qy     1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      783 TGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTA 842

Qy     1050 GluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      843 GAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGAT 902

Qy     1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      903 GCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCT 962

Qy     1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db      963 CACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAATGGAG 1022

Qy     1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1023 TCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAG 1082

Qy     1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1083 GACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACA 1142

Qy     1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1143 GAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGG 1202

Qy     1170 ArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1203 AGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAAT 1262

Qy     1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1263 GGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGAC 1322

Qy     1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1323 ATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGA 1382

Qy     1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1383 AGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGT 1442

Qy     1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1443 GGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACA 1502

Qy     1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1503 GTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGT 1562

Qy     1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1563 GGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGG 1622
```

```
Qy    1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 TTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCC 1682

Qy    1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 TGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCG 1742

Qy    1350 LeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGly 1369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 CTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGG 1802

Qy    1370 LeuLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArgValGlnLysGlnLys 1389
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 CTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAA 1862

Qy    1390 HisLeuProLeuArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHis 1409
           |||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db    1863 CATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGCGGAGAATTTATTCCAT 1922

Qy    1410 GluMetGluThrCysLeuLysArgGluAspProHisGlyThrArgThrSerAspAspThr 1429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 GAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACC 1982

Qy    1430 ProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeuGlyValLeuProAlaSer 1449
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 CCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCT 2042

Qy    1450 GluAlaThrLys 1453
           ||||||||||||
Db    2043 GAAGCCACAAAA 2054
```

RESULT 4
HSM130AC1

```
LOCUS       HSM130AC1             3786 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen cytoplasmic variant 1.
ACCESSION   Z22969
VERSION     Z22969.1  GI:312143
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3786)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE   2  (bases 1 to 3786)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
FEATURES             Location/Qualifiers
     source          1..3786
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3557
                     /codon_start=1
                     /product="M130 antigen cytoplasmic variant 1"
                     /protein_id="CAA80542.1"
                     /db_xref="GI:312144"
                     /db_xref="SPTREMBL:Q07899"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSRYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVA
                     LSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQS
                     QTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICD
                     DSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGW
                     GQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVG
                     VVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSE
                     ETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALK
                     AFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDISVQK
                     TPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQ
                     IQYREMNSCLNADDLDLMNSSENSHESADFSAAELISVSKFLPISGMEKEAILSHTEK
                     ENGNL"
     sig_peptide     102..221
     mat_peptide     222..3554
                     /product="M130 antigen cytoplasmic variant 1"
                     /evidence=experimental
     polyA_signal    3760..3765
     polyA_signal    3764..3769
     polyA_signal    3768..3773
     polyA_site      3783
BASE COUNT     1006 a    746 c   1074 g    960 t
ORIGIN

Alignment Scores:
Pred. No.:              2.77e-215     Length:         3786
Score:                  3441.50       Matches:        608
Percent Similarity:     70.24%        Conservative:   166
Best Local Similarity:  55.17%        Mismatches:     311
Query Match:            42.29%        Indels:         17
DB:                     9             Gaps:           6

US-09-759-130B-381 (1-1453) x HSM130AC1 (1-3786)

Qy     364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
           |||   |||  ||||||||||||   ||||||    |||   |||||||||||||||||
Db     222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy     384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
           |||:::::   |||:::|||   |||:::|||:::            |||      |||
Db     282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy     404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
              |||:::|||   ||||||||||||||   ::::::                  |||
Db     342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy     424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
              |||        :::        ||||||::::::   :::||||||  .||||||||||||
Db     402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy     444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
           |||||||||||||||   :::||||||       |||   :::   |||       :::
Db     462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy     463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
           ||||||||||||      |||||||||      ::::::|||::::::||||||
Db     522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581

Qy     483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
           :::      |||   ||||||:::|||:::|||:::||||||   |||||||||||||||
Db     582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641

Qy     503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
           |||   ::::::      :::|||:::|||:::|||:::||||||   ||||||
Db     642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy     523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
           :::         |||   :::   |||   |||   |||||||||||||||   ||||||
Db     702 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761
```

```
Qy    543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db    762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy    563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db    822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy    583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db    882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy    603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db    942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy    623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db   1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy    643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db   1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy    663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db   1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy    683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db   1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy    703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db   1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy    723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db   1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy    743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db   1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy    763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db   1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy    783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db   1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy    803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db   1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy    823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db   1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy    843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db   1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy    863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db   1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy    883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db   1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841

Qy    903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db   1842 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1901

Qy    923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db   1902 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1961

Qy    943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db   1962 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 2021

Qy    963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db   2022 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 2081

Qy    983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db   2082 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 2141

Qy   1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
Db   2142 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATT 2201

Qy   1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db   2202 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 2261

Qy   1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db   2262 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 2321

Qy   1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db   2322 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 2381

Qy   1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
Db   2382 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441

Qy   1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
Db   2442 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 2501

Qy   1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
Db   2502 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 2561

Qy   1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
Db   2562 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGA 2621

Qy   1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
Db   2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 2681

Qy   1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
Db   2682 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 2741

Qy   1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
Db   2742 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 2801

Qy   1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
Db   2802 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 2861

Qy   1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db   2862 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 2921

Qy   1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
```

```
Db    2922 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981

Qy    1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
           ||||||||||||||||||||||    ||||||   |||    :::::::|||   ||||||
Db    2982 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 3041

Qy    1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
           ||||||||||||   |||||||||:::::::::::::||||||||||||||||||    |||
Db    3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 3101

Qy    1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
           |||||||||   |||::::    ||||||   |||:::||||||||||||||||||||||
Db    3102 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 3161

Qy    1343 ValArgCysSerGlyGlnSerLeuLys-----------------SerLeuAsnAlaSer 1356
           |||   |||:::      |||:::::::                      :::          |||
Db    3162 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 3221

Qy    1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
           |||         :::    |||            |||   |||::::::||||||    :::|||
Db    3222 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 3281

Qy    1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
           :::          |||||||||          |||   |||:::|||:::          |||
Db    3282 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 3332

Qy    1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
           ||||||:::|||                 :::    :::       :::    ||||||     :::|||
Db    3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 3392

Qy    1415 LeuLysArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCys 1434
           |||       :::|||                  :::|||::::::        :::|||
Db    3393 CTGAATGCAGATGATCTGGACCTAATGAATTCCTCAGAAAAT------TCCCATGAGTCA 3446

Qy    1435 GluAspAlaSerAspThrSerLeuLeuGlyVal---------LeuProAlaSerGluAla 1451
              |||    |||            |||:::    |||                ||||||    |||
Db    3447 GCTGATTTCAGTGCTGCTGAACTAATTTCTGTGTCTAAATTTCTTCCTATTTCTGGAATG 3506

Qy    1452 ThrLys 1453
              |||
Db    3507 GAAAAG 3512

RESULT 5
HSM130A
LOCUS       HSM130A                  3703 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen.
ACCESSION   Z22968
VERSION     Z22968.1  GI:312141
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3703)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
REFERENCE   2  (bases 1 to 3703)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
FEATURES             Location/Qualifiers
     source          1. .3703
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102. .3452
                     /codon_start=1
                     /product="M130 antigen"
                     /protein_id="CAA80541.1"
                     /db_xref="GI:312142"
                     /db_xref="SPTREMBL:Q07898"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSRYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVA
                     LSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQS
                     QTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICD
                     DSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGW
                     GQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVG
                     VVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSE
                     ETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALK
                     AFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDISVQK
                     TPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQ
                     IQYREMNSCLNADDLDLMNSSGGHSEPH"
     sig_peptide     102. .221
     mat_peptide     222. .3449
                     /product="M130 antigen"
                     /evidence=experimental
     polyA_signal    3677. .3682
     polyA_signal    3681. .3686
     polyA_signal    3685. .3690
     polyA_site      3700
BASE COUNT      983 a    731 c   1059 g    930 t
ORIGIN

Alignment Scores:
Pred. No.:                1.76e-214      Length:          3703
Score:                    3429.00        Matches:         598
Percent Similarity:       71.27%         Conservative:    161
Best Local Similarity:    56.15%         Mismatches:      294
Query Match:              42.14%         Indels:          12
DB:                       9              Gaps:            4

US-09-759-130B-381 (1-1453) x HSM130A (1-3703)

Qy     364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
           |||   |||   ||||||||||||   ||||||   |||   ||||||||||||||||||
Db     222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy     384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
           |||::::::   |||:::|||   |||:::|||:::          |||       |||
Db     282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy     404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
              |||:::|||   |||||||||||||||   ::::::               |||
Db     342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy     424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
              |||       :::          ||||||::::::   :::||||||   ||||||||||||
Db     402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy     444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
           |||||||||||||||   :::||||||      |||   :::   |||        :::
Db     462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy     463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
           ||||||||||||   |||||||||   ::::::|||::::::||||||
Db     522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581

Qy     483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
           :::   |||   ||||||:::|||:::|||:::||||||   |||||||||||||||
Db     582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641
```

```
Qy    503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db    642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy    523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db    702 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761

Qy    543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db    762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy    563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db    822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy    583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db    882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy    603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db    942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy    623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db   1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy    643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db   1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy    663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db   1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy    683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db   1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy    703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db   1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy    723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db   1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy    743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db   1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy    763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db   1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy    783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db   1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy    803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db   1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy    823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db   1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy    843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db   1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy    863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db   1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy    883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db   1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841

Qy    903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db   1842 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1901

Qy    923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db   1902 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1961

Qy    943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db   1962 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 2021

Qy    963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db   2022 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 2081

Qy    983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db   2082 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 2141

Qy   1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
Db   2142 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATT 2201

Qy   1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db   2202 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 2261

Qy   1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db   2262 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 2321

Qy   1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db   2322 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 2381

Qy   1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
Db   2382 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441

Qy   1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
Db   2442 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 2501

Qy   1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
Db   2502 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 2561

Qy   1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
Db   2562 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGA 2621

Qy   1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
Db   2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 2681

Qy   1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
Db   2682 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 2741

Qy   1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
Db   2742 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 2801

Qy   1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
```

```
             |||:::|||||||||:::|||::::::||||||:::||||||||||||||||||||
Db     2802 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 2861

Qy     1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
             ::::::::||||||::::::    |||     |||    |||||||||||||||||||||||
Db     2862 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 2921

Qy     1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
             |||    ||||||||||||||||||||||||||||||||||||||||      :::|||:::|||
Db     2922 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981

Qy     1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
             |||||||||||||||||||||||     ||||||     |||     ::::::|||     ||||||
Db     2982 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 3041

Qy     1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
             ||||||||||||     |||||||||:::::::::::|||||||||||||||||||     |||
Db     3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 3101

Qy     1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
             |||||||||     |||:::     ||||||     |||:::||||||||||||||||||||||
Db     3102 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 3161

Qy     1343 ValArgCysSerGlyGlnSerLeuLys------------------SerLeuAsnAlaSer 1356
             |||     |||:::     |||::::::                    :::          |||
Db     3162 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 3221

Qy     1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
             |||          :::     |||          |||     |||::::::||||||     :::|||
Db     3222 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 3281

Qy     1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
             :::           |||||||||          |||     |||:::|||:::          |||
Db     3282 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 3332

Qy     1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
             ||||||:::|||                :::     :::          :::     ||||||     :::|||
Db     3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 3392

Qy     1415 LeuLysArgGluAsp 1419
             |||          :::|||
Db     3393 CTGAATGCAGATGAT 3407
```

RESULT 6
HSM130AC2

```
LOCUS       HSM130AC2               4950 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
ACCESSION   Z22970
VERSION     Z22970.1  GI:312145
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4950)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE   2  (bases 1 to 4950)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
FEATURES             Location/Qualifiers
     source          1..4950
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3572
                     /codon_start=1
                     /product="M130 antigen cytoplasmic variant 2"
                     /protein_id="CAA80543.1"
                     /db_xref="GI:312146"
                     /db_xref="SPTREMBL:Q07900"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSRYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVA
                     LSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQS
                     QTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICD
                     DSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGW
                     GQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVG
                     VVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSE
                     ETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALK
                     AFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDISVQK
                     TPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQ
                     IQYREMNSCLNADDLDLMNSSGLWVLGGSIAQGFRSVAAVEAQTFYFDKQLKKSKNVI
                     GSLDAYNGQE"
     sig_peptide     102..221
     mat_peptide     222..3569
                     /product="M130 antigen cytoplasmic variant 2"
                     /evidence=experimental
     polyA_signal    4924..4929
     polyA_signal    4928..4933
     polyA_signal    4932..4937
     polyA_site      4947
BASE COUNT     1348 a    970 c   1277 g   1355 t
ORIGIN

Alignment Scores:
Pred. No.:               2.5e-214      Length:          4950
Score:                   3429.00       Matches:         598
Percent Similarity:      71.27%        Conservative:    161
Best Local Similarity:   56.15%        Mismatches:      294
Query Match:             42.14%        Indels:          12
DB:                      9             Gaps:            4

US-09-759-130B-381 (1-1453) x HSM130AC2 (1-4950)

Qy      364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
             |||   |||   ||||||||||||   |||||        |||   ||||||||||||||||||
Db      222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy      384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
             |||::::::   |||:::|||   |||:::|||:::          |||          |||
Db      282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy      404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
             |||:::|||   ||||||||||||||||   ::::::                        |||
Db      342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy      424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
             |||        :::         ||||||::::::   :::||||||   ||||||||||||
Db      402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy      444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
             |||||||||||||||   :::||||||          |||   :::   |||        :::
Db      462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy      463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
             ||||||||||||   |||||||||   ::::::|||::::::||||||
Db      522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581

Qy      483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
```

```
Db     582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641

Qy     503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db     642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy     523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db     702 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761

Qy     543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db     762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy     563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db     822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy     583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db     882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy     603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db     942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy     623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db    1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy     643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db    1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy     663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db    1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy     683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db    1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy     703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db    1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy     723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db    1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy     743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db    1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy     763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db    1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy     783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db    1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy     803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db    1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy     823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db    1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy     843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db    1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy     863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db    1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy     883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db    1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841

Qy     903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db    1842 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1901

Qy     923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db    1902 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1961

Qy     943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db    1962 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 2021

Qy     963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db    2022 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 2081

Qy     983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db    2082 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 2141

Qy    1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
Db    2142 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGCCCAACAAGGCCTACCATT 2201

Qy    1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db    2202 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 2261

Qy    1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db    2262 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 2321

Qy    1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db    2322 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 2381

Qy    1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
Db    2382 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441

Qy    1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
Db    2442 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 2501

Qy    1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
Db    2502 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 2561

Qy    1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
Db    2562 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGA 2621

Qy    1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
Db    2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 2681

Qy    1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
Db    2682 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 2741

Qy    1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
Db    2742 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 2801
```

```
Qy     1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
               |||:::|||||||||:::|||::::::||||||:::||||||||||||||||||||
Db     2802 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 2861

Qy     1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
            :::::::::||||||::::::   |||   |||   ||||||||||||||||||||||
Db     2862 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 2921

Qy     1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
            |||   ||||||||||||||||||||||||||||||||||||||||    :::|||:::|||
Db     2922 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981

Qy     1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
            ||||||||||||||||||||||||   ||||||   |||   :::::::|||   ||||||
Db     2982 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 3041

Qy     1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
            ||||||||||||   |||||||||::::::::::::||||||||||||||||||   |||
Db     3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 3101

Qy     1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
            |||||||||   |||:::   ||||||   |||:::||||||||||||||||||||||
Db     3102 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 3161

Qy     1343 ValArgCysSerGlyGlnSerLeuLys------------------SerLeuAsnAlaSer 1356
            |||   |||:::      |||::::::                :::         |||
Db     3162 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 3221

Qy     1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
            |||         :::   |||         |||   |||:::::::||||||   :::|||
Db     3222 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 3281

Qy     1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
            :::         ||||||||||   |||   |||:::|||:::         |||
Db     3282 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 3332

Qy     1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
            ||||||:::|||            :::   :::      :::   ||||||   :::|||
Db     3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 3392

Qy     1415 LeuLysArgGluAsp 1419
            |||      :::|||
Db     3393 CTGAATGCAGATGAT 3407
```

RESULT 7
HSM130AE

```
LOCUS       HSM130AE                3802 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen extracellular variant.
ACCESSION   Z22971
VERSION     Z22971.1  GI:312147
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3802)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE   2  (bases 1 to 3802)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
FEATURES             Location/Qualifiers
     source          1..3802
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3551
                     /codon_start=1
                     /product="M130 antigen extracellular variant"
                     /protein_id="CAA80544.1"
                     /db_xref="GI:312148"
                     /db_xref="SPTREMBL:Q07901"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSSKTQKTSLIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRV
                     ELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTG
                     TEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVA
                     CIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATG
                     SAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTS
                     EASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMSI
                     PMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEI
                     WHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNE
                     SSLWDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVV
                     LLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQIQYREMNSCLNADDLDLMNSSGGHS
                     EPH"
     sig_peptide     102..221
     mat_peptide     222..3548
                     /product="M130 antigen extracellular variant"
                     /evidence=experimental
     polyA_signal    3776..3781
     polyA_signal    3780..3785
     polyA_signal    3784..3789
     polyA_site      3799
BASE COUNT     1009 a    751 c   1081 g    961 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              2.41e-212      Length:          3802
Score:                  3396.50        Matches:         597
Percent Similarity:     69.03%         Conservative:    161
Best Local Similarity:  54.37%         Mismatches:      295
Query Match:            41.74%         Indels:          45
DB:                     9              Gaps:            5
```

US-09-759-130B-381 (1-1453) x HSM130AE (1-3802)

```
Qy      364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
            |||   |||   ||||||||||||   ||||||   |||   ||||||||||||||||||
Db      222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy      384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
            |||::::::   |||:::|||   |||:::|||:::      |||         |||
Db      282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy      404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
            |||:::|||   |||||||||||||||   ::::::               |||
Db      342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy      424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
               |||      :::      ||||||::::::   :::||||||   ||||||||||||
Db      402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy      444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
            ||||||||||||||   :::||||||      |||   :::   |||      :::
Db      462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy      463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
            ||||||||||||   |||||||||   ::::::|||::::::||||||
Db      522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581
```

```
Qy     483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db     582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641

Qy     503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db     642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy     523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db     702 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761

Qy     543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db     762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy     563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db     822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy     583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db     882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy     603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db     942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy     623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db    1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy     643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db    1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy     663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db    1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy     683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db    1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy     703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db    1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy     723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db    1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy     743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db    1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy     763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db    1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy     783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db    1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy     803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db    1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy     823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db    1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy     843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db    1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy     863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db    1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy     883 ThrCysIleHisSerArgGluValGlyValValCysSerArg------------------ 896
Db    1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGTAAGACCCAGAAAACATCT 1841

Qy     896 ------------------------------------------------------------ 896
Db    1842 TTAATTGGTTCTTATACTGTGAAAGGGACAGGGTTAGGGAGTCATAGCTGTCTTTTTCTA 1901

Qy     897 --------------------TyrThrAspValArgLeuValAsnGlyLysSerGlnCys 909
Db    1902 AAGCCCTGTCTCCTTCCAGGATACACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGT 1961

Qy     910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929
Db    1962 GAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCTCTGTAACTCTCACTGG 2021

Qy     930 AspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949
Db    2022 GACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACC 2081

Qy     950 ThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCys 969
Db    2082 CCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAGGCATATGTTTCACTGC 2141

Qy     970 LeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCys 989
Db    2142 ACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCTAGGTGCTTCATTATGT 2201

Qy     990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009
Db    2202 CCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTCCCAAACACTGTCCTCG 2261

Qy    1010 CysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIle 1029
Db    2262 TGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCC 2321

Qy    1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049
Db    2322 TGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTA 2381

Qy    1050 GluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069
Db    2382 GAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAGTGAT 2441

Qy    1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089
Db    2442 GCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGGTTCTGCT 2501

Qy    1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109
Db    2502 CATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAAATGCAATGGAAAAGAA 2561

Qy    1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129
Db    2562 TCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAATTGCAGGCACAAGGAG 2621

Qy    1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149
Db    2622 GATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGACCAGTGAAGCCAGCAGA 2681

Qy    1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169
Db    2682 GAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTGGGGCACTGTTGGCAAG 2741

Qy    1170 ArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
```

```
Db    2742 AGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCAGACAAA 2801

Qy    1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
Db    2802 GGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTGGACAAT 2861

Qy    1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
Db    2862 GTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGGGAGAAG 2921

Qy    1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
Db    2922 AGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAACAAGATAAGACTTCAG 2981

Qy    1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
Db    2982 GAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGAGGTTCCTGGGGACA 3041

Qy    1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
Db    3042 GTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTTGGCTGT 3101

Qy    1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
Db    3102 GGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCGATATGG 3161

Qy    1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
Db    3162 CTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGATTGTCCTGCCAGACGC 3221

Qy    1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
Db    3222 TGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAATTGCACAGATATTTCA 3281

Qy    1350 LeuLys------------------SerLeuAsnAlaSerSerGlyHisLeuAlaLeuIle 1363
Db    3282 GTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCGTCAGTCATCCTTTATT 3341

Qy    1364 LeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIle------LeuPheLeuThr 1381
Db    3342 GCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGCATTATTCTTCTTGACT 3401

Qy    1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGly 1401
Db    3402 AAAAAGCGAAGACAGAGACAGCGG---------CTTGCAGTTTCCTCAAGAGGAGAGAAC 3452

Qy    1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAsp 1419
Db    3453 TTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGATGAT 3506

RESULT 8
AX281768
LOCUS       AX281768               3707 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 177 from Patent WO0177389.
ACCESSION   AX281768
VERSION     AX281768.1  GI:16609019
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
            Mikita,T. and Tai,J.
  TITLE     Genes expressed in foam cell differentiation
  JOURNAL   Patent: WO 0177389-A 177 18-OCT-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3707
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 236574.12"
BASE COUNT      984 a    731 c   1062 g    930 t
ORIGIN

Alignment Scores:
Pred. No.:              4.38e-211      Length:         3707
Score:                  3377.00        Matches:        598
Percent Similarity:     71.00%         Conservative:   161
Best Local Similarity:  55.94%         Mismatches:     294
Query Match:            41.50%         Indels:         16
DB:                     6              Gaps:           4

US-09-759-130B-381 (1-1453) x AX281768 (1-3707)

Qy     364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSer-GlyArgValGl 383
Db     222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGGAGAGTGGA 281

Qy     383 uValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAl 403
Db     282 AGTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGT 341

Qy     403 aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLy 423
Db     342 CTCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAA 401

Qy     423 sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe 443
Db     402 TTCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTC 461

Qy     443 rAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSe 462
Db     462 AGCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACA 521

Qy     462 rAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHi 482
Db     522 AGATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGG 581

Qy     482 sSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHi 502
Db     582 GAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGA 641

Qy     502 sAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPr 522
Db     642 TGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGC 701

Qy     522 oMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAs 542
Db     702 TGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGA 761

Qy     542 pValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLy 562
Db     762 TCTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAA 821

Qy     562 sHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGl 582
Db     822 GCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAG 881

Qy     582 yLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGl 602
Db     882 CCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGG 941

Qy     602 yArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSe 622
Db     942 AGAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAA 1001

Qy     622 rGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTy 642
Db    1002 GCAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATT 1061

Qy     642 rGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCy 662
Db    1062 TGGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATG 1121
```

```
Qy     662 sArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCy 682
           |:::::       |||||    :::   |||:::|||:::||||||    ||||||    ||
Db    1122 TAAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATG 1181

Qy     682 sSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLy 702
           ||||||   ||||||:::||||||||||||   ||||||    |||||||||||||||
Db    1182 TTCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGAC 1241

Qy     702 sValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAs 722
            ||||||||   :::|||      :::|||    :::|||      |||||||||:::
Db    1242 AGTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAA 1301

Qy     722 nIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGl 742
               |||:::|||||||||||||||||     ||||||||||||::::::   |||     ::
Db    1302 AGAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCA 1361

Qy     742 uProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGl 762
           :   :::         :::      :::      :::      |||   |||    |||
Db    1362 AGTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAA 1421

Qy     762 yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs 782
            |||    ||||||||||||||          |||:::|||            |||
Db    1422 TGAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTA 1481

Qy     782 nMetGluAlaSerLeuIleCysSer-AlaHisArg-GlnProArgLeuValGlyAlaAsp 801
               ||||||   :::    ||||||  |||||||||  :::||||||||||||||     |||
Db    1482 TGAAGAAGCCAAAATTACCTGCTCAGGCCCACAGGGGAACCCAGACTGGTTGGAGGGGAC 1541

Qy     802 MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp 821
           :::||||||||||||||||||||||||||||||      |||||||||    |||:::||||||
Db    1542 ATTCCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGAT 1601

Qy     822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla 841
           ||||||||||||||    ||||||:::||||||||||||||||||     ||||||
Db    1602 TCAGACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTT 1661

Qy     842 IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys 861
           :::|||:::      |||    |||||||||:::||||||||      ||||||||:::
Db    1662 GTCTCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAA 1721

Qy     862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881
           ||||||||||||||    |||:::||||||:::||||||||||:::               ||||||
Db    1722 TTCCAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAA 1781

Qy     882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg 901
               ||||||    |||||||||:::||||||||||||||||||||||||||||:::::::|||
Db    1782 GGAACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGC 1841

Qy     902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921
           |||||||||||||||:::     |||:::|||:::||||||:::           ||||||    |||
Db    1842 TTGGTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGG 1901

Qy     922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941
           ||||||||||||:::::::|||||||||    |||||||||    |||||||||:::||||||
Db    1902 GGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTT 1961

Qy     942 SerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArg 961
              ||||||    ||||||||||||            |||:::    |||    :::    :::
Db    1962 AAATGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAG 2021

Qy     962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981
           :::|||    |||    |||||||||    |||    |||      :::     :::|||    :::
Db    2022 ATCTGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTA 2081

Qy     982 ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001
           |||   |||||||||          |||                |||||||||||||:::|||
Db    2082 ACTGCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGA 2141

Qy    1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla 1021
           :::    :::|||    |||        |||    ::::::               |||
Db    2142 AACCAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACC 2201

Qy    1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041
           :::||||||     ||||||:::      |||:::|||         :::||||||||||||:::|||
Db    2202 ATTCCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGA 2261

Qy    1042 AspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCys 1061
                 |||||||||||||||||||||||||||||:::|||    ||||||||||||||
Db    2262 GGAGGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGT 2321

Qy    1062 AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081
           ||||||    ||||||||||||||||||||||||||||||:::::||||||||||||
Db    2322 GATGACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAG 2381

Qy    1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp 1101
           |||   |||||||||   |||||||||||||||||||||:::||||||||||||||||||
Db    2382 GCCATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGAT 2441

Qy    1102 AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly 1121
           ::::::    |||    |||    ||||||    :::|||||||||    |||    ||||||||
Db    2442 GAGATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGG 2501

Qy    1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141
           |||    :::|||||||||||||||||||||||||||||||||||||||||    :::|||
Db    2502 CAGCAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTG 2561

Qy    1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
           ||||||    ||||||    |||:::||||||||||||||||||||||||||||||
Db    2562 AGACTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAAT 2621

Qy    1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181
           |||    ||||||:::||||||:::    :::::::::              |||:::||||||
Db    2622 GGAGCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGC 2681

Qy    1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201
           |||||||||||||    :::    |||    :::::    |||    |||    |||
Db    2682 AGGCAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATG 2741

Qy    1202 SerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp-Gl 1221
           |||      |||||||||||:::::::||||||||||||            :::::::|||  ||
Db    2742 TCCATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGGCA 2801

Qy    1221 nCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleTh 1241
           ||||      |||:::|||||||||:::|||:::::|||||||:::||||||||||||||||
Db    2802 GTGCCCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCAC 2861

Qy    1241 rCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIl 1261
           |||::::::::::||||||::::::    |||    |||    ||||||||||||||||||
Db    2862 ATGTGACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGAT 2921

Qy    1261 eTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGl 1281
           |||||||    ||||||||||||||||||||||||||||||||||||    :::|||::
Db    2922 CTGGCATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCA 2981

Qy    1281 uValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPh 1301
           :||||||||||||||||||||||||||||    ||||||    |||    :::::::|||    ||
Db    2982 GGTGGTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTT 3041

Qy    1301 eGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPh 1321
           |||||||||||||||    |||||||||::::::::::::||||||||||||||||||
Db    3042 TGGTCAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTC 3101

Qy    1321 eLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAl 1341
            ||||||||||||    |||:::    ||||||    |||:::||||||||||||||||||||||
Db    3102 CTTGTGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGC 3161

Qy    1341 aGlyValArgCysSerGlyGlnSerLeuLys-----------------SerLeuAsnAl 1355
           |   |||    |||:::      |||::::::                  :::
Db    3162 TGCAGTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCG 3221

Qy    1355 aSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLe 1375
            ||||||         :::    |||           |||    |||::::::||||||   ::
Db    3222 CTCATCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCAT 3281

Qy    1375 uPheIle------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLe 1393
           :|||:::          |||||||||       |||     |||:::|||:::              ||
```

```
Db    3282 TTTCGTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CT 3332

Qy    1393 uArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluTh 1413
           |    ||||||:::|||              :::   :::    :::   ||||||    ::
Db    3333 TGCAGTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTC 3392

Qy    1413 rCysLeuLysArgGluAsp 1419
           :||||||      :::|||
Db    3393 TTGCCTGAATGCAGATGAT 3411

RESULT 9
AF274883
LOCUS       AF274883                4405 bp    mRNA    linear   ROD 10-MAY-2001
DEFINITION  Mus musculus macrophage hemoglobin scavenger receptor CD163
            precursor, mRNA, complete cds.
ACCESSION   AF274883
VERSION     AF274883.1  GI:13236130
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4405)
  AUTHORS   Schaer,D.J., Boretti,F.S., Hongegger,A., Poehler,D., Linnscheid,P.,
            Staege,H., Muller,C., Schoedon,G. and Schaffner,A.
  TITLE     Molecular cloning and characterization of the mouse CD163
            homologue, a highly glucocorticoid-inducible member of the
            scavenger receptor cysteine-rich family
  JOURNAL   Immunogenetics 53 (2), 170-177 (2001)
  MEDLINE   21243729
   PUBMED   11345593
REFERENCE   2  (bases 1 to 4405)
  AUTHORS   Schaer,D.J., Boretti,F.S. and Schaffner,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2000) Departement of Medicine, Medizinische
            Klinik B, University Hospital Zurich, Ramistrasse 100, Zurich, ZH
            8091, Switzerland
FEATURES             Location/Qualifiers
     source          1..4405
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="6"
                     /map="15.43 centirays distal from the D6Mit300"
     CDS             85..3450
                     /note="FAT; member of the scavenger receptor cysteine rich
                     type B superfamily; transmembrane protein type I"
                     /codon_start=1
                     /product="macrophage hemoglobin scavenger receptor CD163
                     precursor"
                     /protein_id="AAK16065.1"
                     /db_xref="GI:13236131"
                     /translation="MGGHRMVLLGGAGSPGCKRFVHLGFFVVAVSSLLSASAVTNAPG
                     EMKKELRLAGGENNCSGRVELKIHDKWGTVCSNGWSMNEVSVVCQQLGCPTSIKALGW
                     ANSSAGSGYIWMDKVSCTGNESALWDCKHDGWGKHNCTHEKDAGVTCSDGSNLEMRLV
                     NSAGHRCLGRVEIKFQGKWGTVCDDNFSKDHASVICKQLGCGSAISFSGSAKLGAGSG
                     PIWLDDLACNGNESALWDCKHRGWGKHNCDHAEDVGVICLEGADLSLRLVDGVSRCSG
                     RLEVRFQGEWGTVCDDNWDLRDASVVCKQLGCPTAISAIGRVNASEGSGQIWLDNISC
                     EGHEATLWECKHQEWGKHYCHHREDAGVTCSDGADLELRLVGGGSRCAGIVEVEIQKL
                     TGKMCSRGWTLADADVVCRQLGCGSALQTQAKIYSKTGATNTWLFPGSCNGNETTFWQ
                     CKNWQWGGLSCDNFEEAKVTCSGHREPRLVGGEIPCSGRVEVKHGDVWGSVCDFDLSL
                     EAASVVCRELQCGTVVSILGGAHFGEGSGQIWGEEFQCSGDESHLSLCSVAPPLDRTC
                     THSRDVSVVCSRYIDIRLAGGESSCEGRVELKTLGAWGPLCSSHWDMEDAHVLCQQLK
                     CGVAQSIPEGAHFGKGAGQVWSHMFHCTGTEEHIGDCLMTALGAPTCSEGQVASVICS
                     GNQSQTLLPCSSLSPVQTTSSTIPKESEVPCIASGQLRLVGGGGRCAGRVEVCHEGSW
                     GTVCDDNWDMTDANVVCKQLDCGVAINATGSAYFGEGAGAIWLDEVICTGKESHIWQC
                     HSHGWGRHNCRHKEDAGVICSEFMSLRLTNEAHKENCTGRLEVFYNGTWGSIGSSNMS
                     PTTVGVVCRQLGCADNGTVKPIPSDKTPSRPMWVDRVQCPKGVDTLWQCPSSPWKQRQ
                     ASPSSQESWIICDNKIRLQEGHTDCSGRVEIWHKGSWGTVCDDSWDLNDAKVVCKQLG
                     CGQAVKALKEAAFGPGTGPIWLNEIKCRGNESSLWDCPAKPWSHSDCGHKEDASIQCL
                     PKMTSESHHGTGHPTLTALLVCGAILLVLLIVFLLWTLKRRRIQRLTVSSRGEVLIHQ
                     VQYQEMDSKADDLDLLKSSGVIQRHTEKENDNL"
     sig_peptide     85..198
     misc_feature    3220..3288
                     /note="Region: transmembrane domain"
BASE COUNT     1174 a    884 c   1246 g   1101 t
ORIGIN

Alignment Scores:
Pred. No.:              3.85e-204      Length:         4405
Score:                  3272.00        Matches:        582
Percent Similarity:     68.11%         Conservative:   174
Best Local Similarity:  52.43%         Mismatches:     332
Query Match:            40.21%         Indels:         22
DB:                     10             Gaps:           9

US-09-759-130B-381 (1-1453) x AF274883 (1-4405)

Qy     320 GlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrp 339
             |||     :::|||   |||   |||                |||    :::   |||:::
Db      91 GGACACAGAATGGTTCTTCTTGGAGGTGCTGGATCTCCTGGTTGTAAAAGGTTTGTC--- 147

Qy     340 AspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerVal 359
                    |||   |||                                       |||
Db     148 ---------CATCTAGGTTTCTTTGTTGTGGCTGTGAGCTCACTTCTCAGTGCCTCTGCT 198

Qy     360 IleCysSerAspGlyAlaAspLeu------GluLeuArgLeuAlaAspGlySerAsnAsn 377
           :::   :::            ::::::      ||||||||||||||   |||   ||||||
Db     199 GTCACTAACGCTCCTGGAGAAATGAAGAAGGAACTGAGACTGGCGGGTGGTGAAAACAAC 258

Qy     378 CysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsn 397
           |||||||||||||||||||||||||:::::||||||::::::|||   |||:::|||
Db     259 TGTAGTGGGAGAGTGGAACTTAAGATCCATGACAAGTGGGGCACAGTGTGCAGTAACGGC 318

Qy     398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417
           |||         :::      |||||||||:::||||||||||||||||    |||:::
Db     319 TGGAGCATGAATGAAGTGTCCGTGGTTTGCCAGCAGCTGGGATGCCCAACTTCTATTAAA 378

Qy     418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437
                       |||       |||       :::       ||||||::::::      :::|||
Db     379 GCCCTTGGATGGGCTAACTCCAGCGCCGGCTCTGGATATATCTGGATGGACAAAGTTTCT 438

Qy     438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr 457
           ||||||||||||||||||||||||||||||||||||    :::||||||          |||
Db     439 TGTACAGGGAATGAGTCAGCTCTTTGGGACTGCAAACATGATGGGTGGGGAAAGCATAAC 498

Qy     458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477
           |||            ||||||||||||   |||||||||   ::::::|||::::::|||
Db     499 TGTACCCATGAAAAAGATGCTGGAGTGACCTGCTCAGATGGATCTAATTTGGAGATGAGA 558

Qy     478 LeuVal---GlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGlu 496
           ||||||      |||         |||   ||||||:::|||:::|||:::||||||:::
Db     559 CTGGTGAACAGTGCGGGCCACCGATGCTTAGGAAGAGTAGAAATAAAGTTCCAGGGAAAG 618

Qy     497 TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln 516
           |||||||||||||||   |||    :::|||       :::|||:::|||:::|||||||||
Db     619 TGGGGGACGGTGTGTGACGACAACTTCAGCAAAGATCACGCTTCTGTGATTTGTAAACAG 678

Qy     517 LeuGlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGly 536
           ||||||||||||         :::         |||                  ||||||
Db     679 CTTGGATGTGGAAGTGCCATTAGTTTCTCTGGCTCAGCTAAATTGGGAGCTGGTTCTGGA 738

Qy     537 ProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu 556
           ||||||||||||||||||:::::||||   ||||||||||||     :::||||||||||:::
Db     739 CCAATCTGGCTCGATGACCTGGCATGCAATGGAAATGAGTCAGCTCTCTGGGACTGCAAA 798

Qy     557 HisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer 576
           |||   |||||||||||||||||||||   |||   |||||||||   |||   |||
Db     799 CACCGGGGATGGGGCAAGCATAACTGTGACCATGCTGAGGATGTCGGTGTGATTTGCTTA 858

Qy     577 GlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeu 596
                 |||         ||||||||||||   |||   :::||||||||||||||||||
Db     859 GAGGGAGCAGATCTGAGCCTGAGACTAGTGGATGGAGTGTCCAGATGTTCAGGAAGATTG 918

Qy     597 GluValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAla 616
           ||||||   |||||||||   ||||||||||||||||||||   |||:::      :::
Db     919 GAAGTGAGATTCCAAGGAGAATGGGGGACCGTGTGTGATGATAACTGGGATCTCCGGGAT 978
```

```
Qy     617 AlaAlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGly 636
Db     979 GCTTCTGTGGTGTGCAAGCAACTGGGATGTCCAACTGCCATCAGTGCCATTGGTCGAGTT 1038

Qy     637 AsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGlu 656
Db    1039 AATGCCAGTGAGGGATCTGGACAGATTTGGCTTGACAACATTTCATGCGAAGGACATGAG 1098

Qy     657 SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu 676
Db    1099 GCAACTCTTTGGGAGTGTAAACACCAAGAGTGGGGAAAGCATTACTGTCATCATAGAGAA 1158

Qy     677 AspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySer 696
Db    1159 GACGCTGGTGTGACATGTTCTGATGGAGCAGATCTGGAACTTAGACTTGTAGGTGGAGGC 1218

Qy     697 SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla 716
Db    1219 AGTCGCTGTGCTGGCATTGTGGAGGTGGAGATTCAGAAGCTGACTGGGAAGATGTGTAGC 1278

Qy     717 AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla 736
Db    1279 CGAGGCTGGACACTGGCAGATGCGGATGTGGTTTGCAGACAGCTTGGATGTGGATCTGCG 1338

Qy     737 IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn 756
Db    1339 CTTCAAACCCAGGCTAAGATCTACTCTAAAACTGGGGCAACAAATACGTGGCTCTTTCCT 1398

Qy     757 SerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln 776
Db    1399 GGATCTTGTAATGGAAATGAAACTACTTTTTGGCAATGCAAAAACTGGCAGTGGGGCGGC 1458

Qy     777 ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArg 796
Db    1459 CTTTCCTGTGATAATTTCGAAGAAGCCAAAGTTACCTGCTCAGGCCACAGGGAACCCAGA 1518

Qy     797 LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp 816
Db    1519 CTGGTTGGAGGAGAAATCCCATGCTCTGGTCGTGTGGAAGTGAAACACGGAGACGTGTGG 1578

Qy     817 ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu 836
Db    1579 GGCTCCGTCTGTGATTTTGACTTGTCTCTGGAAGCTGCCAGTGTGGTGTGCAGGGAATTA 1638

Qy     837 AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu 856
Db    1639 CAATGTGGAACAGTCGTCTCTATCCTAGGGGGAGCACATTTTGGAGAAGGAAGTGGACAG 1698

Qy     857 ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle 876
Db    1699 ATCTGGGGTGAAGAATTCCAGTGTAGTGGGGATGAGTCCCATCTTTCACTATGCTCAGTG 1758

Qy     877 ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg 896
Db    1759 GCGCCCCCGCTAGACAGAACTTGTACCCACAGCAGGGATGTCAGCGTAGTCTGCTCACGA 1818

Qy     897 TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
Db    1819 TACATAGATATTCGTCTGGCAGGCGGCGAGTCCTCCTGTGAGGGAAGAGTGGAGCTCAAG 1878

Qy     917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
Db    1879 ACACTCGGAGCCTGGGGTCCCCTCTGCAGTTCTCATTGGGACATGGAAGATGCTCATGTC 1938

Qy     937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
Db    1939 TTATGTCAGCAGCTGAAGTGTGGGGTTGCCCAATCTATTCCAGAAGGAGCACATTTTGGG 1998

Qy     957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
Db    1999 AAAGGAGCTGGTCAGGTCTGGAGTCACATGTTCCACTGCACTGGAACTGAGGAACATATA 2058

Qy     977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
Db    2059 GGAGATTGCCTCATGACTGCTCTGGGTGCGCCGACGTGTTCCGAAGGACAGGTGGCCTCT 2118

Qy     997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCys------LeuAlaAsnVal 1014
Db    2119 GTCATCTGCTCAGGAAACCAATCCCAGACACTATTGCCATGTAGTTCATTGTCTCCAGTC 2178

Qy    1015 SerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLys 1034
Db    2179 CAAACAACA---AGCTCTACAATTCCAAAGGAGAGTGAAGTTCCCTGCATAGCAAGTGGC 2235

Qy    1035 ArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAsp 1054
Db    2236 CAGCTTCGCTTGGTAGGTGGAGGTGGTCGCTGCGCTGGAAGAGTGGAGGTCTGCCACGAG 2295

Qy    1055 GlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCys 1074
Db    2296 GGCTCTTGGGGCACCGTCTGTGATGACAATTGGGATATGACTGATGCCAATGTGGTGTGC 2355

Qy    1075 GlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGly 1094
Db    2356 AAGCAGCTGGACTGTGGCGTGGCAATTAACGCCACTGGCTCTGCTTACTTCGGGGAAGGA 2415

Qy    1095 SerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGln 1114
Db    2416 GCAGGAGCTATCTGGCTAGACGAAGTCATCTGCACTGGGAAAGAGTCTCATATTTGGCAG 2475

Qy    1115 CysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIle 1134
Db    2476 TGCCATTCACATGGCTGGGGACGCCATAACTGCAGGCACAAAGAAGATGCAGGTGTTATC 2535

Qy    1135 CysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGly 1154
Db    2536 TGCTCCGAGTTCATGTCTCTGAGGCTGACCAACGAAGCCCACAAAGAAAACTGCACAGGT 2595

Qy    1155 ArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThr 1174
Db    2596 CGCCTTGAAGTGTTTTACAATGGTACATGGGGCAGTATTGGCAGTAGCAATATGTCTCCA 2655

Qy    1175 AlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu 1194
Db    2656 ACCACTGTGGGGGTGGTGTGCCGTCAGCTGGGCTGTGCAGACAACGGGACTGTGAAACCC 2715

Qy    1195 AlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLys 1214
Db    2716 ATACCTTCAGACAAGACACCATCCAGGCCCATGTGGGTAGATCGTGTGCAGTGTCCAAAA 2775

Qy    1215 ThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro 1234
Db    2776 GGAGTTGACACTTTGTGGCAGTGCCCCTCGTCACCTTGGAAACAGAGACAGGCCAGCCCC 2835

Qy    1235 ---AlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThr 1253
Db    2836 TCCTCCCAGGAGTCCTGGATCATCTGTGACAACAAAATAAGACTCCAGGAAGGGCATACA 2895

Qy    1254 GluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAsp 1273
Db    2896 GACTGTTCTGGACGTGTGGAGATCTGGCACAAAGGTTCCTGGGGAACAGTGTGTGATGAC 2955

Qy    1274 SerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeu 1293
Db    2956 TCCTGGGATCTTAATGATGCTAAGGTTGTATGTAAGCAGTTGGGCTGTGGCCAAGCTGTG 3015

Qy    1294 AlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMet 1313
Db    3016 AAGGCACTAAAAGAAGCAGCATTTGGTCCAGGAACTGGGCCCATATGGCTCAATGAAATT 3075

Qy    1314 ArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSer 1333
Db    3076 AAGTGTAGAGGGAATGAGTCTTCCCTGTGGGATTGTCCTGCCAAACCGTGGAGTCACAGC 3135

Qy    1334 AspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeu 1353
```

```
                |||||||||||||||||||||||    :::::::|||     :::
Db    3136 GACTGTGGGCACAAAGAAGATGCTTCCATCCAGTGCCTCCCAAAAATGACTTCAGAATCA 3195

Qy    1354 AsnAlaSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeu 1373
              :::      :::||||||      |||               :::    |||    :::||||||
Db    3196 CATCATGGCACAGGTCACCCCACCCTCACGGCACTCTTGGTTTGTGGAGCCATTCTATTG 3255

Qy    1374 ValLeuPheIleLeuPhePheLeuThrTrpCys-----------ArgValGlnLysGlnLys 1389
           ||||||   |||:::||||||    |||                 |||:::|||:::
Db    3256 GTCCTCCTCATTGTCTTCCTCCTGTGGACTCTGAAGCGACGACGGATTCAGCGA------ 3309

Qy    1390 HisLeuProLeuArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHis 1409
                     |||    ||||||:::|||           :::     :::       :::
Db    3310 ---------CTTACAGTTTCCTCAAGAGGAGAGGTCTTGATACATCAAGTTCAGTACCAA 3360

Qy    1410 GluMetGluThrCysLeuLysArgGluAsp 1419
           ||||||::::::         |||    :::|||
Db    3361 GAGATGGATTCA------AAGGCGGATGAT 3384

RESULT 10
BBWC11MR
LOCUS       BBWC11MR                 4475 bp    mRNA    linear   MAM 24-MAY-1994
DEFINITION  B.bovis WC1.1 mRNA.
ACCESSION   X63723
VERSION     X63723.1  GI:13
KEYWORDS    WC1.1 protein.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 4475)
  AUTHORS   Wijngaard,P.L., MacHugh,N.D., Metzelaar,M.J., Romberg,S.,
            Bensaid,A., Pepin,L., Davis,W.C. and Clevers,H.C.
  TITLE     Members of the novel WC1 gene family are differentially expressed
            on subsets of bovine CD4-CD8- gamma delta T lymphocytes
  JOURNAL   J. Immunol. 152 (7), 3476-3482 (1994)
  MEDLINE   94194107
   PUBMED   7511649
REFERENCE   2  (bases 1 to 4475)
  AUTHORS   Wijngaard,P.L.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-1992) P.L.J. Wijngaard, Dept of Immunology,
            University Hospital, P O Box 85500, 3508 GA Utrecht, THE
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .4475
                     /organism="Bos taurus"
                     /strain="indicus"
                     /isolate="D409"
                     /db_xref="taxon:9913"
                     /clone="pBoWC1.1"
                     /cell_line="D409/N2"
                     /tissue_type="blood"
                     /clone_lib="D409/N2.H1"
                     /dev_stage="gamma/delta T-cell"
     gene            22. .4332
                     /gene="WC1.1 subcluster gene"
     CDS             22. .4332
                     /gene="WC1.1 subcluster gene"
                     /codon_start=1
                     /product="BoWC1.1"
                     /protein_id="CAA45255.1"
                     /db_xref="GI:14"
                     /db_xref="SWISS-PROT:P30205"
                     /translation="MALGRHLSLRGLCVLLLGTMVGGQALELRLKDGVHRCEGRVEVK
                     HQGEWGTVDGYRWTLKDASVVCRQLGCGAAIGFPGGAYFGPGLGPIWLLYTSCEGTES
                     TVSDCEHSNIKDYRNDGYNHGRDAGVVCSGFVRLAGGDGPCSGRVEVHSGEAWIPVSD
                     GNFTLATAQIICAELGCGKAVSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPC
                     PGGTCHHSGSAQVVCSAYSEVRLMTNGSSQCEGQVEMNISGQWRALCASHWSLANANV
                     ICRQLGCGVAISTPGGPHLVEEGDQILTARFHCSGAESFLWSCPVTALGGPDCSHGNT
                     ASVICSGNQIQVLPQCNDSVSQPTGSAASEDSAPYCSDSRQLRLVDGGGPCAGRVEIL
                     DQGSWGTICDDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKES
                     HVWRCPSRGWGQHNCRHKQDAGVICSEFLALRMVSEDQQCAGWLEVFYNGTWGSVCRN
                     PMEDITVSTICRQLGCGDSGTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWN
                     YNSCSPKEEAYIWCADSRQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVC
                     KQLGCGEALDATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAG
                     VICSGFVRLAGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLG
                     HMPFRESDGQVWAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMK
                     NGTSQCEGQVEMKISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQ
                     ISTAQFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAG
                     SAASEESSPYCSDSRQLRLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQL
                     GCGEALNATGSAHFGAGSGPIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVIC
                     SEFLALRMVSEDQQCAGWLEVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTS
                     VGLREGSRPRWVDLIQCRKMDTSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTA
                     AACTDREKLRLRGGDSECSGRVEVWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEA
                     VRSAAFGPGNGSIWLDEVQCGGRESSLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLP
                     TTTAGTRTTSNSLPGIFSLPGVLCLILGSLLFLVLVILVTQLLRWRAERRALSSYEDA
                     LAEAVYEELDYLLTQKEGLGSPDQMTDVPDENYDDAEEVPVPGTPSPSQGNEEEVPPE
                     KEDGVRSSQTGSFLNFSREAANPGEGEESFWLLQGKKGDAGYDDVELSALGTSPVTFS
                     "
     polyA_site      4453
BASE COUNT      933 a   1113 c   1442 g    987 t
ORIGIN

Alignment Scores:
Pred. No.:              9.9e-192        Length:         4475
Score:                  3082.00         Matches:        613
Percent Similarity:     55.85%          Conservative:   208
Best Local Similarity:  41.70%          Mismatches:     495
Query Match:            37.87%          Indels:         154
DB:                     4               Gaps:           23

US-09-759-130B-381 (1-1453) x BBWC11MR (1-4475)

Qy      28 CysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu 47
           |||:::||||||||||||                       :::          |||    :||||||
Db      58 TGTGTCCTCCTCCTCGGC---------------ACCATGGTGGGTGGTCAAGCTCTGGAG 102

Qy      48 LeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
           |||||||||   :::|||           |||   |||   ||||||||||||   ||||||
Db     103 CTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCACCAAGGA 162

Qy      68 GlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrValValCysLys 87
           :::||||||||||||                 |||          ::::::|||||||||:::
Db     163 GAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGTGTGCAGA 222

Qy      88 GlnLeuGlyCysProPheSerPheAlaMet---------PheArgPheGlyGlnAlaVal 104
           ||||||||||||        :::                   ||||||      :::
Db     223 CAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTTGGGCCAGGACTT 282

Qy     105 ThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeu 124
                    |||   |||||||||          ||||||   |||   ||||||   :::
Db     283 ---------GGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTC 333

Qy     125 TrpGluCysGlnHis---------ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141
              :::|||:::|||          :::::::::   :::              ||||||
Db     334 AGTGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGG 393

Qy     142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161
           |||   ||||||   |||   |||                  :::||||||      |||:::
Db     394 GATGCTGGAGTAGTCTGCTCAGGA------------TTTGTGCGTCTGGCTGGAGGGGAT 441

Qy     162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
              |||||||||||||||||||||               |||   |||       :::   |||
Db     442 GGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGAT 501

Qy     182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
              :::   |||   ||||||   ::::::|||   :::|||||||||         :::
Db     502 GGGAACTTCACACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCT 561

Qy     202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
                    |||                            :::|||   ::::::
Db     562 GTGTCTGTCCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAG 621

Qy     222 IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241
```

```
                     |||:::|||    |||        ||||||    |||
Db     622 TTCAGGTGTGAGGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGG 681

Qy     242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu 261
                |||   |||:::               :::   |||        |||::::::
Db     682 GGCACGTGTCACCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGAAGTCCGG 741

Qy     262 LeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281
           |||   :::         |||::::::::|||   |||:::||||||:::   |||   |||
Db     742 CTC---ATGACAAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGA 798

Qy     282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLys 301
           :::|||      :::|||            |||:::   |||   |||:::|||:::|||:::
Db     799 CAATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGT 858

Qy     302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321
           ||||||||||||||   |||:::               |||   ||||||||
Db     859 CAGCTCGGCTGTGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGT 918

Qy     322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
           |||   :::                     |||||||||   |||||||||||||||   |||
Db     919 GATCAGATCCTAACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGT 978

Qy     342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
              :::                  ||||||   |||   |||         ||||||||||||
Db     979 CCTGTGACTGCCCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGC 1038

Qy     362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
           |||
Db    1039 TCAGGAAAC--------------------------------------------------- 1047

Qy     382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
              :::::::::::      |||
Db    1048 ---CAGATCCAGGTGCTTCCCCAG------------------------------------ 1068

Qy     402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
                                         |||      ||||||
Db    1069 ------------------------------TGCAACGACTCCGTG--------------- 1083

Qy     422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
           ::::::|||:::      |||
Db    1084 TCTCAACCTACAGGCTCTGCGGCCTCA--------------------------------- 1110

Qy     442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461

Db    1110 ------------------------------------------------------------ 1110

Qy     462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
              |||:::         |||||||||               ||||||||||||
Db    1111 GAGGACAGCGCCCCCTACTGCTCAGACAGCAGG------CAGCTCCGCCTGGTGGACGGG 1164

Qy     482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
                    ||||||   ||||||:::|||:::         ||||||   ||||||||:::|||
Db    1165 GGCGGTCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGT 1224

Qy     502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
              |||   |||         :::|||   |||||||||:::|||||||||||||||:::
Db    1225 GATGACGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAA 1284

Qy     522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
              ::::::      |||      :::|||      |||||||||||||||||||||
Db    1285 GCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGAC 1344

Qy     542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
           :::::::::|||   |||   ||||||::::::|||   |||            ||||||||
Db    1345 AACTTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGG 1404

Qy     562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
           :::|||||||||   |||::::::|||         |||   ||||||
Db    1405 CAGCACAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCA------GAGTTCCTG 1458

Qy     582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601
              ||||||:::|||               :::|||:::|||   |||||||||::::::

Db    1459 GCCCTCAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAAT 1518

Qy     602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCys 621
           |||   |||||:::||||||   :::                     :::   :::|||
Db    1519 GGGACCTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGC 1578

Qy     622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
              ||||||   |||      |||               |||      |||:::|||
Db    1579 AGACAGCTTGGCTGTGGGGACAGT------------GGAACCCTCAACTCTTCTGTTGCT 1626

Qy     641 ---------GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
                    |||:::      |||:::|||   :::   |||         ::::::
Db    1627 CTTAGAGAAGGTTTTAGGCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACC 1686

Qy     658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
              ||||||   |||   :::         |||         |||   ||||||      |||:::
Db    1687 TCTCTCTGGCAGTGTCCTTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAA 1746

Qy     678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697
              :::   |||:::|||:::          :::::::||||||||   |||
Db    1747 GCCTATATCTGGTGTGCAGACAGC------AGACAGATCCGCCTGGTGGATGGAGGTGGT 1800

Qy     698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
           ||||||:::|||:::||||||:::         ||||||:::   |||   :::|||   :::
Db    1801 CGCTGCTCTGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGAC 1860

Qy     718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIle 737
              |||   ::::::   |||   |||||||||:::||||||   ||||||   |||:::
Db    1861 CGCTGGGACCTGGACGATGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTG 1920

Qy     738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
              :::                  |||         :::      |||   :::
Db    1921 GACGCCACTGTCTCTTCCTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTG 1980

Qy     758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
              |||   |||   |||:::   :::|||   |||         |||   |||:::|||
Db    1981 AACTGCAGAGGAGAGGAGTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACAC 2040

Qy     778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
              |||:::            :::|||   :::|||||||||                  ||||||
Db    2041 AACTGCAATCATCAAGAAGATGCAGGAGTCATCTGCTCAGGATTT------GTGCGTCTG 2094

Qy     798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
              |||   |||   ||||||||||||||||||||||||         :::   |||
Db    2095 GCTGGAGGAGATGGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACC 2154

Qy     818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
              |||   |||   :::|||:::|||   |||   |||:::|||   ||||||
Db    2155 CCAGTGTCTGATGGAAACTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGA 2214

Qy     838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
           ||||||   |||:::|||:::               |||   :::   :::|||
Db    2215 TGTGGCAAGGCTGTGTCTGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTC 2274

Qy     858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
           |||||||||:::|||:::|||:::|||   |||         |||      ||||||   |||
Db    2275 TGGGCTGAAGAGTTCAGGTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTG 2334

Qy     878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
              |||      ||||||:::||||||                  ||||||||||||   |||
Db    2335 CCCTGTCCAGGAGGCACATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTAC 2394

Qy     898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
           |||:::|||:::|||:::   ||||||   |||||||||:::|||||||||||||||:::
Db    2395 ACAGAAGTCCAGCTTATGAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAG 2454

Qy     917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
           :::   |||   |||   :::||||||   :::||||||         :::|||   |||
Db    2455 ATCTCTGGACGATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTT 2514

Qy     937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
           :::||||||||||||   ||||||   |||:::||||||         |||   ::::::
Db    2515 GTCTGTCGTCAGCTCGGCTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTG 2574
```

```
Qy     957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
           |||        ::::::          :::|||||||||  |||  ||||||  |||
Db    2575 GAAGGAGGTGATCAGATCTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTG 2634

Qy     977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
              :::|||   :::|||  ||||||  |||  |||  |||||||||||  |||
Db    2635 TGGAGTTGTCCTGTGACTGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCT 2694

Qy     997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
           |||||||||:::|||:::   ||||||  |||     |||        :::|||
Db    2695 GTGATCTGCTCAGGAAACCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAA 2754

Qy    1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
           |||     ||||||     |||   |||:::        |||   :::    ::::::|||
Db    2755 CCTGCAGGCTCTGCGGCCTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTC 2814

Qy    1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
           |||||||||||||||         |||   |||||||||||||||         |||
Db    2815 CGCCTGGTGGACGGGGGCGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCC 2874

Qy    1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
           ||||||||||||||||||||||   |||||||||   ||||||   |||||||||::::::
Db    2875 TGGGGCACCATCTGTGATGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAG 2934

Qy    1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
           ||||||||||||   |||   |||||||||   |||||||||||||||   |||||||||
Db    2935 CTGGGCTGTGGAGAAGCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGG 2994

Qy    1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
           |||||||||||||||||||||||||||||||||   |||||||||:::|||:::|||||||
Db    2995 CCCATCTGGCTGGACGACCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCT 3054

Qy    1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
           ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db    3055 TCCCGGGGCTGGGGGCGGCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCA 3114

Qy    1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
           ||||||   |||||||||:::   |||         |||   :::   |||||||||  |||
Db    3115 GAGTTCCTGGCCCTCAGGATGGTGAGC------GAGGACCAGCAGTGTGCTGGGTGGCTG 3168

Qy    1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
           |||||||||||||||||||||||||||||||||||   |||         :::
Db    3169 GAGGTTTTCTACAACGGGACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACT 3228

Qy    1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
              ::::|||||||||||||||||||||||:::::::|||   ::::::     :::
Db    3229 GTGTCCGTGATCTGCAGACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTT 3288

Qy    1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
                 :::   ||||||          |||||||||  |||||||||  |||
Db    3289 GGTCTCAGGGAAGGTTCTAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGAT 3348

Qy    1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
              |||:::|||||||||   |||   ||||||:::            ||||||   |||
Db    3349 ACCTCTCTCTGGCAGTGTCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAG 3408

Qy    1237 GluThrTrpIleThrCysGlu-------------------------------------- 1243
           |||   :::|||:::||||||
Db    3409 GAAGCCTACATCTCATGTGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGC 3468

Qy    1244 ---------AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
                    ::::::|||:::|||||||||:::||||||||||||||||||||||||
Db    3469 ACAGACAGAGAGAAGCTCCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAG 3528

Qy    1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
           :::||||||   |||||||||||||||||||||||||||||||||   ||||||||||||
Db    3529 GTGTGGCACAACGGCTCCTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCT 3588

Qy    1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
           ||||||||||||||||||||||||||||   ||||||   |||:::|||   |||:::
Db    3589 GAGGTGGTGTGTCAGCAGCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCA 3648

Qy    1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
           ||||||   |||   |||:::|||||||||||||||:::::::::|||  |||  ||||||
Db    3649 TTTGGCCCTGGAAATGGGAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCC 3708

Qy    1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
           ||||||||||||   |||:::||||||||||||||||||||||||   |||:::||||||
Db    3709 TCCCTGTGGGACTGTGTTGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGAT 3768

Qy    1341 AlaGlyValArgCysSerGly-------------------------------------- 1347
           |||||||||||||||||||||
Db    3769 GCTGGTGTGAGGTGCTCTGGTGTAAGGACAACATTGCCCACGACCACAGCAGGGACCAGA 3828

Qy    1348 ---------GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeu 1364
                    ||||||      :::   :::         |||   |||   |||||||||
Db    3829 ACAACCTCAAATTCTCTCCCTGGCATCTTCTCCCTGCCTGGGGTTCTCTGCCTTATCCTG 3888

Qy    1365 SerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg 1384
              |||:::   |||:::|||:::::::|||          |||   |||   |||
Db    3889 GGGTCGCTTCTCTTCCTGGTCCTCGTCATCCTGGTGACTCAGCTACTCAGATGG---AGA 3945

Qy    1385 ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu 1404
              ::::::                   |||      :::         :::|||
Db    3946 GCAGAGCGC------------------AGAGCCTTATCCAGCTATGAAGATGCTCTTGCT 3987

Qy    1405 GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg 1424
           |||   ::::::   |||::::::      |||   :::::::::      |||:::
Db    3988 GAAGCTGTGTATGAGGAGCTCGATTACCTTCTGACACAGAAGGAAGGTCTGGGCAGCCCA 4047

Qy    1425 ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu 1443
                    |||   |||:::            :::||||||   :::      :::
Db    4048 GATCAGATGACTGATGTCCCTGATGAAAATTATGATGATGCTGAAGAAGTACCAGTGCCT 4107

Qy    1444 GlyValLeuProAlaSerGluAlaThrLys 1453
           |||               |||:::   :::
Db    4108 GGAACTCCTTCTCCCTCTCAGGGGAATGAG 4137
```

```
RESULT 11
HSA243224
LOCUS       HSA243224               7656 bp    mRNA    linear   PRI 27-NOV-2000
DEFINITION  Homo sapiens mRNA for DMBT1 protein 8kb transcript variant 1
            (DMBT1/8kb.1).
ACCESSION   AJ243224
VERSION     AJ243224.1  GI:6624921
KEYWORDS    DMBT1 gene; DMBT1/8kb.1 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7656)
  AUTHORS   Mollenhauer,J., Holmskov,U., Wiemann,S., Krebs,I., Herbertz,S.,
            Madsen,J., Kioschis,P., Coy,J.F. and Poustka,A.
  TITLE     The genomic structure of the DMBT1 gene: evidence for a region with
            susceptibility to genomic instability
  JOURNAL   Oncogene 18 (46), 6233-6240 (1999)
  MEDLINE   20065089
   PUBMED   10597221
REFERENCE   2  (bases 1 to 7656)
  AUTHORS   Mollenhauer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-1999) Mollenhauer J., Molecular Genome Analysis,
            Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120
            Heidelberg, GERMANY
FEATURES             Location/Qualifiers
     source          1..7656
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q25.3-q26.1"
                     /tissue_type="lung"
                     /dev_stage="adult"
     gene            1..7318
                     /gene="DMBT1"
     5'UTR           1..106
```

```
                  /gene="DMBT1"
     CDS          107..7318
                  /gene="DMBT1"
                  /function="putative tumour suppressor, role in immune
                  system and differentiation"
                  /note="8kb protein, variant 1
                  alternative splice form of the DMBT1 gene that is
                  expressed in human adult lung"
                  /codon_start=1
                  /product="DMBT1/8kb.1 protein"
                  /protein_id="CAB63942.1"
                  /db_xref="GI:6624922"
                  /db_xref="SPTREMBL:Q9UGM2"
                  /translation="MGISTVILEMCLLWGQVLSTGGWIPRTTDYASLIPSEVPLDTTV
                  AEGSPFPSELTLESTAAEGSPISLESTLESTVAEGSLIPSESTLESTVAEGSDSGLAL
                  RLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNAWFGQG
                  SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLRPESWPVR
                  ISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLG
                  CGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICS
                  APLSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSW
                  DTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSH
                  NCQHSEDAGVICSDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYRGSWGTV
                  CDDSWDTNDANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHN
                  GWLSHNCGHSEDAGVICSGPESSLALGLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTN
                  DANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPNNGWLSHNCG
                  HHEDAGVICSAAQSRSTPRPDTLSTITLPPSTVGSESSLTLRLVNGSDRCQGRVEVLY
                  RGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESY
                  LWSCPHNGWLSHNCGHHEDAGVICSVSQSRPTPSPDTWPTSHASTAGSESSLALRLVN
                  GGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRRLGCGWATSAPGNARFGQGSGPI
                  VLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAHSWSTPSPDTLPTITLPA
                  STVGSESSLALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVVCRQLGCGWAM
                  SAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAGVICSASQSR
                  PTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDA
                  NVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHDGWLSHNCGHH
                  EDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGPW
                  GTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSC
                  PHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNGGDR
                  CQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGSARFGQGSGPIALDD
                  VRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSRASTAGSE
                  STLALRLVNGGDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNA
                  QFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQSTPRPD
                  TWLTTNLPALTVGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDSWDTNDANVVC
                  RQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLTHNCGHHEDAG
                  VICSATQINSTTTDWWHPTTTTTARPSSNCGGFLFYASGTFSSPSYPAYYPNNAKCVW
                  EIEVNSGYRINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICNDTRQIFTSSY
                  NRMTIHFRSDISFQNTGFLAWYNSFPSDATLRLVNLNSSYGLCAGRVEIYHGGTWGTV
                  CDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVECSGTESTLWQCRNR
                  GWFSHNCNHREDAGVICSGNHLSTPAPFLNITRPNTDYSCGGFLSQPSGDFSSPFYPG
                  NYPNNAKCVWDIEVQNNYRVTVIFRDVQLEGGCNYDYIEVFDGPYRSSPLIARVCDGA
                  RGSFTSSSNFMSIRFISDHSITRRGFRAEYYSSPSNDSTNLLCLPNHMQASVSRSYLQ
                  SLGFSASDLVISTWNGYYECRPQITPNLVIFTIPYSGCGTFKQADNDTIDYSNFLTAA
                  VSGGIIKRRTDLRIHVSCRMLQNTWVDTMYIANDTIHVANNTIQVEEVQYGNFDVNIS
                  FYTSSSFLYPVTSRPYYVDLNQDLYVQAEILHSDAVLTLFVDTCVASPYSNDFTSLTY
                  DLIRSGCVRDDTYGPYSSPSLRIARFRFRAFHFLNRFPSVYLRCKMVVCRAYDPSSRC
                  YRGCVLRSKRDVGSYQEKVDVVLGPIQLQTPPRREEEPR"
     mat_peptide  183..7315
                  /gene="DMBT1"
                  /product="DMBT1/8kb.1 protein"
BASE COUNT     1604 a   2126 c   2166 g   1760 t
ORIGIN

Alignment Scores:
Pred. No.:              1.92e-185      Length:          7656
Score:                  2990.00        Matches:         619
Percent Similarity:     52.03%         Conservative:    177
Best Local Similarity:  40.46%         Mismatches:      504
Query Match:            36.74%         Indels:          230
DB:                     9              Gaps:            23

US-09-759-130B-381 (1-1453) x HSA243224 (1-7656)

Qy     44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
             :::   |||  |||||||||||||||||         |||  |||   |||||||||
Db    794 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGGTC 853

Qy     64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThr 83
             :::::|||   ||||||||||||||||||||   |||:::|||         :::
Db    854 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 913

Qy     84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet---------PheArgPhe 100
          |||||||||:::|||||||||||    ::::::    :::             :::|||
Db    914 GTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTT 973

Qy    101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
          ||||||            |||   |||   ||||||||||||   |||   |||:::
Db    974 GGCCAGGGC---------TCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACAC 1024

Qy    121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
          ||||||   ||||||   |||   |||      |||   :::|||||||||   |||
Db   1025 GAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGT 1084

Qy    141 GluAspValGlyValAsnCysTyrGly--------------------------------- 149
          ||||||   ||||||   |||
Db   1085 GAAGACGCTGGTGTCATCTGCTCAGCTCCCCTGTCCCGGCCGACACCCAGCCCAGATACT 1144

Qy    150 ---------------------------------GluAlaAsnLeuGlyLeuArgLeuVal 158
                                           |||::::::|||   ||||||||||||
Db   1145 TGGCCGACCTCACATGCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTG 1204

Qy    159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
          ::::||   :::   |||   |||||||||||||||   ::::::      |||||||||
Db   1205 AATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACC 1264

Qy    179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
          :::|||||||||   |||:::   :::   |||   ||||||||||||||||||||||||
Db   1265 GTGTGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGT 1324

Qy    199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
             :::      :::   |||                              ||||||
Db   1325 GGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTC 1384

Qy    219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
          |||||||||:::   |||   |||   |||      ||||||:::|||   |||   |||
Db   1385 CTGGATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGC 1444

Qy    239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSer--- 257
          |||   :::|||:::|||   |||:::||||||      :::   |||   |||:::
Db   1445 TGGCTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGACACGTTG 1504

Qy    258 ------------------------------------SerAspLeuGluLeuArgLeu 264
                                              |||   |||   |||||||||
Db   1505 CCGACCATCACCTTACCTGCATCGACAGTAGGATCTGAATCCAGTTTGGCCCTGAGGCTG 1564

Qy    265 ValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGly 284
          |||   |||   :::|||||||   ||||||||||||:::      :::|||   ||||||
Db   1565 GTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGC 1624

Qy    285 ThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGly 304
          |||||||||         |||:::         |||:::|||||||||:::|||||||||
Db   1625 ACCGTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGC 1684

Qy    305 CysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
          ||||||   |||:::            |||               ||||||      :::
Db   1685 TGTGGCTGGGCCATGTTGGCCCCAGGAAATGCCCGGTTTGGTCAGGGCTCAGGACCCATT 1744

Qy    325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
             ||||||   |||   |||||||||||||||:::||||||   |||   |||:::
Db   1745 GTCCTGGATGACGTGCGCTGCTCAGGGAATGAGTCCTACTTGTGGAGCTGCCCCCACAAT 1804

Qy    345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer---Asp 363
          |||   ::::::   :::|||   |||      |||      ||||||||||||
Db   1805 GGCTGGCTCTCCCATAACTGTGGCCATAGTGAAGACGCTGGTGTCATCTGCTCAGGACCT 1864

Qy    364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
             :::   |||   |||   |||   :::|||   :::   |||   ||||||||||||
Db   1865 GAATCCAGTTTGGCCCTGGGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAG 1924
```

```
Qy     384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db    1925 GTCCTATACCGAGGCTCTTGGGGCACCGTGTGTGATGACAGCTGGGACACCAATGATGCC 1984

Qy     404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db    1985 AATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGG 2044

Qy     424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db    2045 TTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCC 2104

Qy     444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArgSerAsp 463
Db    2105 TACCTGTGGAGCTGCCCCAACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGAT 2164

Qy     464 AlaGlyValIleCysSer------------------------------------------ 469
Db    2165 GCTGGTGTCATCTGCTCAGCTGCCCAGTCCCGGTCGACGCCCAGGCCAGACACGTTGTCG 2224

Qy     470 ------------------------------AspLysAlaAspLeuAspLeuArgLeuVal 479
Db    2225 ACCATCACGTTACCTCCATCGACAGTAGGATCTGAATCCAGTTTGACCCTGAGGCTGGTG 2284

Qy     480 GlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThr 499
Db    2285 AATGGAAGTGACAGGTGTCAGGGCCGAGTAGAGGTCCTATACCGAGGCTCCTGGGGCACC 2344

Qy     500 ValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCys 519
Db    2345 GTGTGTGATGACAGCTGGGATACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGT 2404

Qy     520 GlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrp 539
Db    2405 GGCTGGGCCACGTCGGCCCCAGGAAATGCCCGGTTTGGCCAGGGCTCAGGACCCATTGTT 2464

Qy     540 LeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGly 559
Db    2465 CTGGATGATGTGCGCTGCTCAGGACACGAGTCCTACCTGTGGAGCTGCCCCCACAATGGC 2524

Qy     560 TrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer--------- 576
Db    2525 TGGCTCTCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGTTTCCCAG 2584

Qy     577 ---------------GlyAspAlaThrTrp------------------------------ 581
Db    2585 TCCCGGCCGACACCCAGTCCAGATACTTGGCCGACCTCACATGCATCAACAGCAGGATCT 2644

Qy     582 ------------GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGlu 597
Db    2645 GAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAG 2704

Qy     598 ValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAla 617
Db    2705 GTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGATAGCTGGGACACCAGTGACGCC 2764

Qy     618 AlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsn 637
Db    2765 AATGTGGTCTGCCGGCGGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGG 2824

Qy     638 AlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
Db    2825 TTTGGCCAGGGTTCAGGACCCATTGTCCTGGATGACGTGCGCTGCTCAGGCTATGAGTCC 2884

Qy     658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
Db    2885 TACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGTCAGCACAGTGAAGAC 2944

Qy     678 ValGlyValIleCysSerAspAla------------------------------------ 685
Db    2945 GCTGGTGTCATCTGCTCAGCTGCCCACTCCTGGTCGACGCCCAGTCCAGACACATTGCCG 3004

Qy     686 ------------------------------------SerAspMetGluLeuArgLeuVal 693
Db    3005 ACCATCACCTTGCCTGCATCGACAGTAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTG 3064

Qy     694 GlyGlySerSerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIle 713
Db    3065 AATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACC 3124

Qy     714 LeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCys 733
Db    3125 GTGTGCGATGACAGCTGGGACACCAATGATGCCAATGTCGTCTGCAGGCAACTGGGCTGT 3184

Qy     734 GlySerAlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeu 753
Db    3185 GGCTGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGTCAGGGCTCAGGACCCATTGTC 3244

Qy     754 MetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGlu 773
Db    3245 CTGGATGATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGCCCCCACAATGGC 3304

Qy     774 TrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArg 793
Db    3305 TGGCTCTCCCACAACTGTGGCCATAGTGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAG 3364

Qy     794 ---GlnPro--------------------------------------------------- 795
Db    3365 TCCCGGCCAACACCTAGTCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCT 3424

Qy     796 -----------------ArgLeuValGlyAlaAspMetProCysSerGlyArgValGlu 809
Db    3425 GAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAG 3484

Qy     810 ValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAla 829
Db    3485 GTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCC 3544

Qy     830 AsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHis 849
Db    3545 AATGTGGTTTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCGG 3604

Qy     850 PheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGluGlySerGluThr 869
Db    3605 TTTGGCCAGGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCC 3664

Qy     870 HisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIleHisSerArgGlu 889
Db    3665 TATCTGTGGAGTTGCCCCCACGATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGAC 3724

Qy     890 ValGlyValValCysSerArgTyrThrAsp------------------------------ 899
Db    3725 GCTGGTGTCATCTGCTCAGCTTCCCAGTCCCAGCCGACACCCAGCCCAGACACTTGGCCA 3784

Qy     900 ------------------------------------------ValArgLeuValAsnGly 905
Db    3785 ACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGA 3844

Qy     906 LysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCys 925
Db    3845 GGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCCCCTGGGGCACCGTGTGT 3904

Qy     926 AspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThr 945
Db    3905 GATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGCTGG 3964

Qy     946 AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHis 965
Db    3965 GCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTGGAT 4024

Qy     966 ArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGly 985
Db    4025 GATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGGCTC 4084

Qy     986 AlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGln 1005
```

```
                :::        |||   |||                   |||||||||:::    |||    :::|||
Db    4085 TCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCAG 4144

Qy    1006 ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGly 1025
           |||       |||                              :::   |||
Db    4145 CCGACA---CCCAGCCCAGACACTTGGCCAACCTCGCATGCATCAACAGCAGGATCTGAA 4201

Qy    1026 SerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCys 1045
           |||:::|||                            ||||||||||||:::|||   ||||||
Db    4202 TCCAGTTTG-------------------GCCCTGAGGCTGGTGAATGGAGGTGACAGGTGT 4243

Qy    1046 AlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrp 1065
              ||||||||||||:::   :::   |||   |||||||||:::|||||||||   |||
Db    4244 CAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGG 4303

Qy    1066 AspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAla 1085
           |||   :::||||||:::||||||||||:::::|||||||||||||   |||   :::|||
Db    4304 GACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCC 4363

Qy    1086 ThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCys 1105
                 ||||||   ||||||:::|||||||||||||||   |||||||||:::    |||
Db    4364 CCAGGAAGTGCCCGGTTTGGCCAGGGTTCAGGACCCATTGCCCTGGATGATGTGCGCTGC 4423

Qy    1106 ThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCys 1125
           :::|||   ||||||:::||||||   ||||||          ||||||     |||:::|||
Db    4424 TCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGT 4483

Qy    1126 ArgHisLysGluAspAlaGlyValIleCys------------------------------ 1135
              |||   ||||||||||||||||||||
Db    4484 GGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCAGCCGACACCCAGC 4543

Qy    1136 ---------------------------------------SerGluPheThr---AlaLeu 1141
                                                   ||||||   |||    ||||||
Db    4544 CCAGACACTTGGCCAACCTCTCGTGCATCAACAGCAGGATCTGAATCCACTTTGGCCCTG 4603

Qy    1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
           ||||||    :::                 :::    |||    ||||||:::||||||   |||
Db    4604 AGACTGGTGAAT------GGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCAA 4657

Qy    1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181
           |||:::||||||:::|||                   |||      |||    :::||||||
Db    4658 GGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAACGTGGTCTGC 4717

Qy    1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLys------ 1199
           ||||||||||||||            |||          ||||||     :::
Db    4718 AGGCAGCTGGGCTGT---------GGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTT 4768

Qy    1200 ---ThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSer 1218
                 |||||||||   :::   :::||||||:::::|||
Db    4769 GGCCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCTTAC 4828

Qy    1219 IleTrpGlnCysLeuSerAlaProTrp--------------------------------- 1227
           :::|||   |||            |||
Db    4829 CTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGATGCT 4888

Qy    1228 -----------------GluArgArgIleSerSerProAlaGluGluThrTrpIleThr 1241
                             :::      |||:::|||      :::|||||||:::|||
Db    4889 GGTGTCATCTGCTCAGCTGCTCAGTCCCAGTCAACGCCCAGGCCAGATACTTGGCTGACC 4948

Qy    1242 --------------------------CysGluAspArgIleArgValArg--------- 1249
                                     |||          :::   :::|||
Db    4949 ACCAACTTACCGGCATTGACAGTAGGATCTGAATCCAGTTTGGCTCTGAGGCTGGTGAAT 5008

Qy    1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
           ||||||||         |||   ||||||||||||:::   :::   |||||||||||||||
Db    5009 GGAGGTGAC---AGGTGTCGAGGCCGAGTGGAGGTCCTGTATCGAGGCTCCTGGGGAACC 5065

Qy    1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
           |||||||||||||||||||          :::|||   |||||||||:::||||||||||||
Db    5066 GTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGT 5125

Qy    1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
           |||   |||:::::|||         :::|||   ||||||||||||:::|||    |||
Db    5126 GGCTGGGCCATGTCGGCCCCAGGAAATGCCCGGTTTGGCCAGGGCTCAGGACCCATTGTC 5185

Qy    1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
           |||||||||:::||||||    ||||||||||||:::||||||    |||        |||
Db    5186 CTGGATGATGTGCGCTGCTCAGGGAATGAGTCCTACCTGTGGAGCTGCCCCCACAAAGGC 5245

Qy    1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
           |||           :::|||||||||    |||||||||||||||    ||||||
Db    5246 TGGCTCACCCACAACTGTGGCCATCACGAAGACGCTGGTGTCATCTGCTCAGCCACCCAA 5305

Qy    1350 LeuLysSerLeuAsnAlaSerSerGlyHis 1359
           :::    |||                |||
Db    5306 ATAAATTCTACTACGACAGATTGGTGGCAT 5335
```

```
RESULT 12
HSA243212
LOCUS       HSA243212               7683 bp    mRNA    linear   PRI 17-SEP-1999
DEFINITION  Homo sapiens mRNA for DMBT1 protein 8kb transcript variant 2
            (DMBT1/8kb.2).
ACCESSION   AJ243212
VERSION     AJ243212.1  GI:5912463
KEYWORDS    DMBT1 gene; DMBT1/8kb.2 protein.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7683)
  AUTHORS   Holmskov,U., Mollenhauer,J., Madsen,J., Vitved,L., Gronlund,J.,
            Tornoe,I., Kliem,A., Reid,K.B., Poustka,A. and Skjodt,K.
  TITLE     Cloning of gp-340, a putative opsonin receptor for lung surfactant
            protein D
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (19), 10794-10799 (1999)
  MEDLINE   99415938
   PUBMED   10485905
REFERENCE   2  (bases 1 to 7683)
  AUTHORS   Mollenhauer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-1999) Mollenhauer J., Molecular Genome Analysis,
            Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120
            Heidelberg, GERMANY
FEATURES             Location/Qualifiers
     source          1..7683
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q25.3-q26.1"
                     /tissue_type="trachea"
                     /dev_stage="adult"
     gene            1..7683
                     /gene="DMBT1"
     5'UTR           1..106
                     /gene="DMBT1"
     CDS             107..7345
                     /gene="DMBT1"
                     /function="putative tumour suppressor; role in immune
                     system and differentiation"
                     /note="Sequence is an alternative splice form of the DMBT1
                     gene that is expressed in human adult trachea. Isoforms of
                     DMBT1 are identical to the collectin binding protein
                     gp-340. Full-length cDNA clone contains 1 bp deletions in
                     codons 100 and 1751, that were corrected by comparison
                     with the genomic exons"
                     /codon_start=1
                     /product="DMBT1/8kb.2 protein"
                     /protein_id="CAB56155.1"
                     /db_xref="GI:5912464"
                     /translation="MGISTVILEMCLLWGQVLSTGGWIPRTTDYASLIPSEVPLDTTV
                     AEGSPFPSELTLESTVAEGSPISLESTLETTVAEGSLIPSESTLESTVAEGSDSGLAL
                     RLVNGDGRCQGRVEILYRGSWGAVCDDSWDTNDANVVCRQLGCGWAMSAPGNAWFGQG
                     SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLRPESWPVR
                     ISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLG
                     CGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICS
                     APQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSW
```

```
                      DTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSH
                      NCQHSEDAGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVE
                      VLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGN
                      ESYLWSCPHNGWLSHNCGHSEDAGVICSGPESSLALRLVNGGDRCQGRVEVLYRGSWG
                      TVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCP
                      NNGWLSHNCGHHEDAGVICSAAQSRSTPRPDTLSTITLPPSTVGSESSLTLRLVNGSD
                      RCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLD
                      DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSVSQSRPTPSPDTWPTSHASTAGP
                      ESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGCGWATSAPGN
                      ARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAHSWSTPSP
                      DTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVV
                      CRQLGCGWAMSAPGNARFGQGSGPIVLDDARCSGHESYLWSCPHNGWLSHNCGHSEDA
                      GVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTV
                      CDDYWDTNDANVACRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHN
                      GWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQG
                      RVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRC
                      SGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSL
                      ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFG
                      QGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAFQSQPTPSPDTWP
                      TSRASTAGSESTLALRLVNGGDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGC
                      GWAMSAPGNAQFGQGSGPIVLDDVRCSGHEPYLWSCPHNGWLSHNCGHHEDAGVICSA
                      AQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDS
                      WDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLGDVRCSGNESYLWSCPHKGWLT
                      HNCGHHEDAGVICSATQINSTTTDWWHPTTTTTARPSSNCGGFLFYASGTFSSPSYPA
                      YYPNNAKCVWEIEVNSGYRINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICN
                      DTRQIFTSSYNRMTIHFRSDISFQNTGFLAWYNSFPSDATLRLVNLNSSYGLCAGRVE
                      IYHGGTWGAVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVECSGT
                      ESTLWQCRNRGWFSHNCNHREDAGVICSGNHLSTPAPFLNITRPNNYSCGGFLSQPSG
                      DFSSPFYPGNYPNNAKCVWDIEVQNNYRVTVIFRDVQLEGGCNYDYIEVFDGPYRSSP
                      LIARVCDGARGSFTSSSNFMSIRFISDHSITRRGFRAEYYSSPSNDSTNLLCLPNHMQ
                      ASVSRSYLQSLGFSASDLVISTWNGYYECRPQITPNLVIFTIPYSGCGTFKQADNDTI
                      DYSNLLTAAVSGGIIKRRTDLRIHVSCRMLQNTWVDTMYIANDTIHVANNTIQVEEVQ
                      YGNFDVNISFYTSSSFLYPVTSRPYYVDLNQDLYVQAEILHSDAVLTLFVDTCVASPY
                      SNDFTSLTYDLIRSGCVRDDTYGPYSSPSLRIARFRFRAFHFLNRFPSVYLRCKMVVC
                      RAYDPSSRCYRGCVLRSKRDVGSYQEKVDVVLGPIQLQTPPRREEEPR"
     sig_peptide      107..184
                      /gene="DMBT1"
     mat_peptide      182..7342
                      /gene="DMBT1"
                      /product="DMBT1/8kb.2 protein"
     3'UTR            7343..7683
                      /gene="DMBT1"
BASE COUNT     1610 a   2141 c   2170 g   1762 t
ORIGIN

Alignment Scores:
Pred. No.:              3.02e-185       Length:         7683
Score:                  2987.00         Matches:        617
Percent Similarity:     51.56%          Conservative:   177
Best Local Similarity:  40.06%          Mismatches:     506
Query Match:            36.70%          Indels:         240
DB:                     9               Gaps:           22

US-09-759-130B-381 (1-1453) x HSA243212 (1-7683)

Qy     44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
Db    794 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGGGTGGAGGTC 853

Qy     64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThr 83
Db    854 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 913

Qy     84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet---------PheArgPhe 100
Db    914 GTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTT 973

Qy    101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
Db    974 GGCCAGGGC---------TCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACAC 1024

Qy    121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140

Db   1025 GAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGT 1084

Qy    141 GluAspValGlyValAsnCysTyrGly--------------------------------- 149
Db   1085 GAAGACGCTGGTGTCATCTGCTCAGCTCCCCAGTCCCGGCCGACACCCAGCCCAGATACT 1144

Qy    150 ---------------------------------GluAlaAsnLeuGlyLeuArgLeuVal 158
Db   1145 TGGCCGACCTCACATGCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTG 1204

Qy    159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
Db   1205 AATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACC 1264

Qy    179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
Db   1265 GTGTGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGT 1324

Qy    199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
Db   1325 GGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTC 1384

Qy    219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
Db   1385 CTGGATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGC 1444

Qy    239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyr--------- 255
Db   1445 TGGCTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCCCAC 1504

Qy    255 ------------------------------------------------------------ 255
Db   1505 TCCTGGTCGACGCCCAGTCCAGACACGTTGCCGACCATCACCTTACCTGCATCGACAGTA 1564

Qy    256 ---AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArg 274
Db   1565 GGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGA 1624

Qy    275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAla 294
Db   1625 GTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGACACCAAT 1684

Qy    295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314
Db   1685 GATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTTGGCCCCAGGAAAT 1744

Qy    315 ProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334
Db   1745 GCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGACGTGCGCTGCTCAGGGAAT 1804

Qy    335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354
Db   1805 GAGTCCTACTTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGTGGCCATAGT 1864

Qy    355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373
Db   1865 GAAGACGCTGGTGTCATCTGCTCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 1924

Qy    374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIle 393
Db   1925 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCTTGGGGCACCGTG 1984

Qy    394 CysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysPro 413
Db   1985 TGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGC 2044

Qy    414 PheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIle 433
Db   2045 TGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTG 2104

Qy    434 AsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLys 453
Db   2105 GATGATGTGCGCTGCTCAGGACATGAGTCCTACCTGTGGAGCTGCCCCAACAATGGCTGG 2164
```

```
Qy    454 AlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSer----------- 469
Db   2165 CTCTCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCCCAGTCC 2224

Qy    469 ------------------------------------------------------------ 469
Db   2225 CGGTCGACGCCCAGGCCAGACACGTTGTCGACCATCACGTTACCTCCATCGACAGTAGGA 2284

Qy    470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489
Db   2285 TCTGAATCCAGTTTGACCCTGAGGCTGGTGAATGGAAGTGACAGGTGTCAGGGCCGAGTA 2344

Qy    490 GluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509
Db   2345 GAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGATACCAATGAT 2404

Qy    510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529
Db   2405 GCCAATGTAGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCGGCCCCAGGAAATGCC 2464

Qy    530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGlu 549
Db   2465 CGGTTTGGCCAGGGCTCAGGACCCATTGTTCTGGATGATGTGCGCTGCTCAGGACACGAG 2524

Qy    550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569
Db   2525 TCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAA 2584

Qy    570 AspValIleValThrCysSer-----------------------GlyAspAlaThrTrp 581
Db   2585 GATGCTGGTGTCATCTGCTCAGTTTCCCAGTCCCGGCCGACACCCAGTCCAGATACTTGG 2644

Qy    582 ----------------------------------------GlyLeuArgLeuValGly 587
Db   2645 CCGACCTCACATGCATCAACAGCAGGACCTGAATCCAGCTTGGCCCTGAGGCTGGTGAAT 2704

Qy    588 GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal 607
Db   2705 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTG 2764

Qy    608 CysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAspCysPro 627
Db   2765 TGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGTGGC 2824

Qy    628 SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu 647
Db   2825 TGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTG 2884

Qy    648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667
Db   2885 GATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGG 2944

Qy    668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla------ 685
Db   2945 CTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCCCACTCC 3004

Qy    685 ------------------------------------------------------------ 685
Db   3005 TGGTCGACGCCCAGTCCAGACACATTGCCGACCATCACCTTGCCTGCATCGACAGTAGGA 3064

Qy    686 ------SerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysVal 703
Db   3065 TCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTG 3124

Qy    704 GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle 723
Db   3125 GAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGCGATGACAGCTGGGACACCAATGAT 3184

Qy    724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743
Db   3185 GCCAATGTCGTCTGCAGGCAACTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCC 3244

Qy    744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763
Db   3245 CGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGATGCGCGCTGCTCAGGACACGAG 3304

Qy    764 AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783
Db   3305 TCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATAGTGAA 3364

Qy    784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro-------------------- 795
Db   3365 GACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCGGCCAACACCTAGTCCAGACACTTGG 3424

Qy    796 ---------------------------------------------ArgLeuValGly 799
Db   3425 CCAACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 3484

Qy    800 AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal 819
Db   3485 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTG 3544

Qy    820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839
Db   3545 TGTGATGACTACTGGGACACCAATGATGCCAATGTGGCTTGCAGGCAGCTGGGCTGTGGC 3604

Qy    840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859
Db   3605 TGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTG 3664

Qy    860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879
Db   3665 GATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGG 3724

Qy    880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899
Db   3725 CTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCC 3784

Qy    899 ------------------------------------------------------------ 899
Db   3785 CAGCCGACACCCAGCCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCTGAA 3844

Qy    900 ------------ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915
Db   3845 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTC 3904

Qy    916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935
Db   3905 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 3964

Qy    936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle 955
Db   3965 GTGGTTTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTT 4024

Qy    956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975
Db   4025 GGCCAGGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTAT 4084

Qy    976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrVal 995
Db   4085 CTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCT 4144

Qy    996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015
Db   4145 GGTGTCATCTGCTCAGCTTCCCAGTCCCAGCCGACA---CCCAGCCCAGACACTTGGCCA 4201

Qy   1016 AspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg 1035
Db   4202 ACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTG-----------------GCC 4243

Qy   1036 LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055
Db   4244 CTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGC 4303

Qy   1056 PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075
```

```
Db   4304 TCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGG 4363

Qy   1076 LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095
Db   4364 CAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCA 4423

Qy   1096 GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115
Db   4424 GGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGC 4483

Qy   1116 ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCys 1135
Db   4484 CCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGC 4543

Qy   1136 SerGluPhe--------------------------------------------------- 1138
Db   4544 TCAGCTTTCCAGTCCCAGCCGACACCCAGCCCAGACACTTGGCCAACCTCTCGTGCATCA 4603

Qy   1139 --------------------ThrAlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151
Db   4604 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGGTGAAT------GGAGGTGACAGG 4657

Qy   1152 CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn 1171
Db   4658 TGTCGAGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGTGATGACTAC 4717

Qy   1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191
Db   4718 TGGGACACCAATGATGCCAACGTGGTCTGCAGGCAGCTGGGCTGT---------GGCTGG 4768

Qy   1192 ValSerLeuAlaProLeuSerLys---------ThrGlySerGlyPheMetTrpValAsp 1208
Db   4769 GCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGAT 4828

Qy   1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227
Db   4829 GATGTGCGCTGCTCAGGACACGAGCCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTC 4888

Qy   1228 ------------------------------------------------GluArgArgIle 1231
Db   4889 TCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCTCAGTCCCAG 4948

Qy   1232 SerSerProAlaGluGluThrTrpIleThr---------------------------Cys 1242
Db   4949 TCAACGCCCAGGCCAGATACTTGGCTGACCACCAACTTACCGGCATTGACAGTAGGATCT 5008

Qy   1243 GluAspArgIleArgValArg---------GlyGlyAspThrGluCysSerGlyArgVal 1259
Db   5009 GAATCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTGTCGAGGCCGAGTG 5065

Qy   1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279
Db   5066 GAGGTCCTGTATCGAGGCTCCTGGGGAACCGTGTGTGATGACAGCTGGGACACCAATGAT 5125

Qy   1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAla 1299
Db   5126 GCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCGGCCCCAGGAAATGCC 5185

Qy   1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
Db   5186 CGGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGGTGATGTGCGCTGCTCAGGGAATGAG 5245

Qy   1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
Db   5246 TCCTACCTGTGGAGCTGCCCCCACAAAGGCTGGCTCACCCACAACTGTGGCCATCACGAA 5305

Qy   1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
Db   5306 GACGCTGGTGTCATCTGCTCAGCCACCCAAATAAATTCTACTACGACAGATTGGTGGCAT 5365

RESULT 13
HSA297935
```

```
LOCUS       HSA297935               7667 bp    mRNA    linear   PRI 12-JAN-2002
DEFINITION  Homo sapiens mRNA for DMBT1 8 kb transcript variant 2 (DMBT1 gene).
ACCESSION   AJ297935
VERSION     AJ297935.1  GI:14715230
KEYWORDS    alternative splicing; deleted in malignant brain tumors 1; DMBT1
            gene; DMBT1/8kb.2 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mollenhauer,J., Herbertz,S., Helmke,B., Kollender,G., Krebs,I.,
            Madsen,J., Holmskov,U., Sorger,K., Schmitt,L., Wiemann,S.,
            Otto,H.F., Grone,H.J. and Poustka,A.
  TITLE     Deleted in Malignant Brain Tumors 1 is a versatile mucin-like
            molecule likely to play a differential role in digestive tract
            cancer
  JOURNAL   Cancer Res. 61 (24), 8880-8886 (2001)
  MEDLINE   21621923
   PUBMED   11751412
REFERENCE   2  (bases 1 to 7667)
  AUTHORS   Mollenhauer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-2000) Mollenhauer J., Molecular Genome Analysis,
            Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120
            Heidelberg, GERMANY
FEATURES             Location/Qualifiers
     source          1..7667
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q25.3-q26.1"
                     /tissue_type="small intestine"
                     /dev_stage="adult"
     gene            88..7329
                     /gene="DMBT1"
     CDS             88..7329
                     /gene="DMBT1"
                     /function="putative role in mucosal protection, immune
                     response, wound healing"
                     /note="DMBT1 8kb transcript variant 2
                     deleted in malignant brain tumors 1
                     alternative splice variant, contains all exons except for
                     exon 17"
                     /codon_start=1
                     /evidence=experimental
                     /product="DMBT1/8kb.2 protein"
                     /protein_id="CAC44122.1"
                     /db_xref="GI:14715231"
                     /translation="MGISTVILEVCLLWGQVLSTGGWIPRTTDYASLIPSEVPLDPTV
                     AEGSPFPSESTLESTVAEGSPISLESTLESTVAEGSLIPSESTLESTVAEGSDSGLAL
                     RLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNAWFGQG
                     SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLRPESWPVR
                     ISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLG
                     CGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICS
                     APQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSW
                     DTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSH
                     NCQHSEDAGVICSAAHSRSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVE
                     VLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGN
                     ESYLWSCPHNGWLSHNCGHSEDAGVICSGPESSLALRLVNGGDRCQGRVEVLYRGSWG
                     TVCDDSWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCP
                     NNGWLSHNCGHHEDAGVICSAAQSRSTPGPDTLSTITLPPSTVGSESSLTLRLVNGSD
                     RCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLD
                     DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSVSQSRPTPSPDTWPTSHASTAGS
                     ESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGCGWATSAPGN
                     ARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAHSWSTPSP
                     DTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVV
                     CRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDA
                     GVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTV
                     CDDYWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHN
                     GWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQG
                     RVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCSWATSAPGNARFGQGSGPIVLDDVRC
                     SGHESYLWSCPHNGWFSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSL
```

```
                      ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFG
                      QGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWP
                      TSRASTAGSESTLALRLVNGGDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGC
                      GWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSA
                      AQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDS
                      WDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLT
                      HNCGHHEDAGVICSATQINSTTTDWWHPTTTTARPSSNCGGFLFYASGTFSSPSYPA
                      YYPNNAKCVWEIEVNSGYRINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICN
                      DTRQIFTSSYNRMTIHFRSDISFQNTGFLAWYNSFPSDATLRLVNLNSSYGLCAGRVE
                      IYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVECSGT
                      ESTLWQCRNRGWFSHNCNHREDAGVICSGNHLSTPAPFLNITRPNTDYSCGGFLSQPS
                      GDFSSPFYPGNYPNNAKCVWDIEVQNNYRVTVIFRDVQLEGGCNYDYIEVFDGPYRSS
                      PLIARVCDGARGSFTSSSNFMSIRFISDHSITRRGFRAEYYSSPSNDSTNLLCLPNHM
                      QASVSRSYLQSLGFSASDLVISTWNGYYECRPQITPNLVIFTIPYSGCGTFKQADNDT
                      IDYSNFLTAAVSGGIIKRRTDLRIHVSCRMLQNTWVDTMYIANDTIHVANNTIQVEEV
                      QYGNFDVNISFYTSSSFLYPVTSRPYYVDLNQDLYVQAEILHSDAVLTLFVDTCVASP
                      YSNDFTSLTYDLIRSGCVRDDTYGPYSSPSLRIARFRFRAFHFLNRFPSVYLRCKMVV
                      CRAYDPSSRCYRGCVLRSKRDVGSYQEKVDVVLGPIQLQTAPRREEEPR"
     sig_peptide      88. .162
                      /gene="DMBT1"
     mat_peptide      163. .7326
                      /gene="DMBT1"
                      /product="DMBT1/8kb.2 protein"
BASE COUNT     1605 a   2137 c   2166 g   1759 t
ORIGIN

Alignment Scores:
Pred. No.:              6.39e-185       Length:         7667
Score:                  2982.00         Matches:        618
Percent Similarity:     51.56%          Conservative:   176
Best Local Similarity:  40.13%          Mismatches:     506
Query Match:            36.64%          Indels:         240
DB:                     9               Gaps:           23

US-09-759-130B-381 (1-1453) x HSA297935 (1-7667)

Qy       44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
             :::   |||   |||||||||||||||||         |||   |||   |||||||||
Db      775 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGGTC 834

Qy       64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThr 83
                :::::|||   ||||||||||||||||||      |||:::|||         :::
Db      835 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 894

Qy       84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet---------PheArgPhe 100
            ||||||||:::||||||||||||   ::::::   :::               :::|||
Db      895 GTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTT 954

Qy      101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
            ||||||                  |||   |||   |||||||||||   |||   |||:::
Db      955 GGCCAGGGC---------TCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACAC 1005

Qy      121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
             ||||||   ||||||   |||   |||      |||   :::||||||||   |||
Db     1006 GAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGT 1065

Qy      141 GluAspValGlyValAsnCysTyrGly--------------------------------- 149
            ||||||   ||||||   |||
Db     1066 GAAGACGCTGGTGTCATCTGCTCAGCTCCCCAGTCCCGGCCGACACCCAGCCCAGATACT 1125

Qy      150 ---------------------------------GluAlaAsnLeuGlyLeuArgLeuVal 158
                                             |||:::::|||   ||||||||||||
Db     1126 TGGCCGACCTCACATGCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTG 1185

Qy      159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
            :::|||   :::   |||   |||||||||||||||   ::::::      |||||||||
Db     1186 AATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACC 1245

Qy      179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
            :::||||||||||   |||:::   :::   |||   ||||||||||||||||||||||||
Db     1246 GTGTGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGT 1305

Qy      199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
                 :::      :::   |||                                    ||||||
Db     1306 GGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTC 1365

Qy      219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
            |||||||||:::   |||   |||   |||      ||||||:::|||   |||   |||
Db     1366 CTGGATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGC 1425

Qy      239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyr--------- 255
            |||   :::|||:::|||   |||:::||||||         :::   |||
Db     1426 TGGCTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCCCAC 1485

Qy      255 ------------------------------------------------------------ 255

Db     1486 TCCAGGTCGACGCCCAGTCCAGACACGTTGCCGACCATCACCTTACCTGCATCGACAGTA 1545

Qy      256 ---AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArg 274
                   |||   |||   |||||||||||||   |||   :::|||||||   ||||||
Db     1546 GGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGA 1605

Qy      275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAla 294
            ||||||:::         :::|||   ||||||||||||||||         |||:::
Db     1606 GTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGACACCAAT 1665

Qy      295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314
               |||:::|||||||||:::||||||||||||||||   |||:::         |||
Db     1666 GATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTTGGCCCCAGGAAAT 1725

Qy      315 ProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334
                           ||||||      :::   ||||||   |||   ||||||||||||
Db     1726 GCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGACGTGCGCTGCTCAGGGAAT 1785

Qy      335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354
            ||||||:::||||||   |||   |||:::|||   ::::::   :::|||   |||
Db     1786 GAGTCCTACTTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGTGGCCATAGT 1845

Qy      355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373
               |||      ||||||||||||         :::   |||   |||||||||   :::
Db     1846 GAAGACGCTGGTGTCATCTGCTCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 1905

Qy      374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIle 393
            |||   :::   |||   |||||||||||||||               |||   |||:::
Db     1906 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCTTGGGGCACCGTG 1965

Qy      394 CysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysPro 413
            ||||||   :::|||               |||   ||||||||:::||||||||||||
Db     1966 TGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGC 2025

Qy      414 PheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIle 433
            ::::::                     |||:::               :::   |||   :::
Db     2026 TGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTG 2085

Qy      434 AsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLys 453
            :::   :::   |||:::|||:::||||||   ||||||   |||      :::|||
Db     2086 GACGATGTGCGCTGCTCAGGACATGAGTCCTACCTGTGGAGCTGCCCCAACAATGGCTGG 2145

Qy      454 AlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSer------------ 469
                              |||            ||||||||||||||||||
Db     2146 CTCTCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCCCAGTCC 2205

Qy      469 ------------------------------------------------------------ 469

Db     2206 CGGTCGACGCCCGGGCCAGACACGTTGTCGACCATCACGTTACCTCCATCGACAGTAGGA 2265

Qy      470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489
                ::::::   |||   ||||||||||||                  |||   ||||||:::
Db     2266 TCTGAATCCAGTTTGACCCTGAGGCTGGTGAATGGAAGTGACAGGTGTCAGGGCCGAGTA 2325

Qy      490 GluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509
            ||||||   |||:::|||   |||||||||||||||   |||   |||   |||   :::
Db     2326 GAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGATACCAATGAT 2385

Qy      510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529
            |||   |||||||||:::||||||||||||||                  |||
```

```
Db   2386 GCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCGGCCCCAGGAAATGCC 2445

Qy    530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGlu 549
Db   2446 CGGTTTGGCCAGGGCTCAGGACCCATTGTTCTGGATGATGTGCGCTGCTCAGGACACGAG 2505

Qy    550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569
Db   2506 TCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAA 2565

Qy    570 AspValIleValThrCysSer---------------------GlyAspAlaThrTrp 581
Db   2566 GATGCTGGTGTCATCTGCTCAGTTTCCCAGTCCCGGCCGACACCCAGTCCAGATACTTGG 2625

Qy    582 -----------------------------------------GlyLeuArgLeuValGly 587
Db   2626 CCGACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 2685

Qy    588 GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal 607
Db   2686 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTG 2745

Qy    608 CysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAspCysPro 627
Db   2746 TGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGTGGC 2805

Qy    628 SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu 647
Db   2806 TGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTG 2865

Qy    648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667
Db   2866 GATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGG 2925

Qy    668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla------ 685
Db   2926 CTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCCCACTCC 2985

Qy    685 ------------------------------------------------------------ 685
Db   2986 TGGTCGACGCCCAGTCCAGACACATTGCCGACCATCACCTTGCCTGCATCGACAGTAGGA 3045

Qy    686 ------SerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysVal 703
Db   3046 TCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTG 3105

Qy    704 GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle 723
Db   3106 GAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGCGATGACAGCTGGGACACCAATGAT 3165

Qy    724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743
Db   3166 GCCAATGTCGTCTGCAGGCAACTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCC 3225

Qy    744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763
Db   3226 CGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAG 3285

Qy    764 AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783
Db   3286 TCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATAGTGAA 3345

Qy    784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro--------------------- 795
Db   3346 GACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCGGCCAACACCTAGTCCAGACACTTGG 3405

Qy    796 ---------------------------------------------ArgLeuValGly 799
Db   3406 CCAACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 3465

Qy    800 AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal 819
Db   3466 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTG 3525

Qy    820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839
Db   3526 TGTGATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGC 3585

Qy    840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859
Db   3586 TGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTG 3645

Qy    860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879
Db   3646 GATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGG 3705

Qy    880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899
Db   3706 CTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCC 3765

Qy    899 ------------------------------------------------------------ 899
Db   3766 CAGCCGACACCCAGCCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCTGAA 3825

Qy    900 ------------ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915
Db   3826 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTC 3885

Qy    916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935
Db   3886 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 3945

Qy    936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle 955
Db   3946 GTGGTTTGCAGGCAGCTGGGCTGTAGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTT 4005

Qy    956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975
Db   4006 GGCCAGGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTAT 4065

Qy    976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrVal 995
Db   4066 CTGTGGAGCTGCCCCCACAATGGCTGGTTCTCCCACAACTGTGGCCATCATGAAGACGCT 4125

Qy    996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015
Db   4126 GGTGTCATCTGCTCAGCTTCCCAGTCCCAGCCGACA---CCCAGCCCAGACACTTGGCCA 4182

Qy   1016 AspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg 1035
Db   4183 ACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTG-----------------GCC 4224

Qy   1036 LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055
Db   4225 CTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGC 4284

Qy   1056 PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075
Db   4285 TCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGG 4344

Qy   1076 LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095
Db   4345 CAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCA 4404

Qy   1096 GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115
Db   4405 GGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGC 4464

Qy   1116 ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCys 1135
Db   4465 CCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGC 4524

Qy   1135 ------------------------------------------------------------ 1135
Db   4525 TCAGCTTCCCAGTCCCAGCCGACACCCAGCCCAGACACTTGGCCAACCTCTCGTGCATCA 4584
```

```
Qy   1136 ---------SerGluPheThr---AlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151
                   |||||    |||   |||||||||||     :::            :::
Db   4585 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGGTGAAT------GGAGGTGACAGG 4638

Qy   1152 CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn 1171
          |||   |||||:::||||||   |||   |||:::||||||:::|||
Db   4639 TGTCGAGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGTGATGACTAC 4698

Qy   1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191
                |||      ||| . :::||||||||||||||||||              |||
Db   4699 TGGGACACCAATGATGCCAACGTGGTCTGCAGGCAGCTGGGCTGT---------GGCTGG 4749

Qy   1192 ValSerLeuAlaProLeuSerLys---------ThrGlySerGlyPheMetTrpValAsp 1208
                ||||||    :::                |||||||||    :::    :::|||
Db   4750 GCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGAT 4809

Qy   1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227
          |||::::::|||                 :::|||   |||           |||
Db   4810 GATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTC 4869

Qy   1228 ---------------------------------------------GluArgArgIle 1231
                                                     .     :::
Db   4870 TCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCTCAGTCCCAG 4929

Qy   1232 SerSerProAlaGluGluThrTrpIleThr-------------------------Cys 1242
          |||:::|||      :::||||||:::|||
Db   4930 TCAACGCCCAGGCCAGATACTTGGCTGACCACCAACTTACCGGCATTGACAGTAGGATCT 4989

Qy   1243 GluAspArgIleArgValArg---------GlyGlyAspThrGluCysSerGlyArgVal 1259
          |||       :::    :::|||         |||||||||      |||   |||||||||
Db   4990 GAATCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTGTCGAGGCCGAGTG 5046

Qy   1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279
          |||:::   :::   |||||||||||||||||||||||||||||||||          :::
Db   5047 GAGGTCCTGTATCGAGGCTCCTGGGGAACCGTGTGTGATGACAGCTGGGACACCAATGAT 5106

Qy   1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAla 1299
          |||   ||||||||:::||||||||||||||||   |||:::::::|||      :::|||
Db   5107 GCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCGGCCCCAGGAAATGCC 5166

Qy   1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
             ||||||||||||:::|||   |||   |||||||||:::||||||   |||||||||
Db   5167 CGGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGGAATGAG 5226

Qy   1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
          |||:::||||||   |||  .   |||   |||          :::|||||||||   |||
Db   5227 TCCTACCTGTGGAGCTGCCCCCACAAAGGCTGGCTCACCCACAACTGTGGCCATCACGAA 5286

Qy   1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
          ||||||||||||   ||||||          :::   |||                  |||
Db   5287 GACGCTGGTGTCATCTGCTCAGCCACCCAAATAAATTCTACTACGACAGATTGGTGGCAT 5346
```

RESULT 14
AF159456
LOCUS   AF159456   7686 bp   mRNA   linear   PRI 16-SEP-1999
DEFINITION Homo sapiens gp-340 variant protein (DMBT1) mRNA, complete cds.
ACCESSION   AF159456
VERSION   AF159456.1 GI:5733597
KEYWORDS   .
SOURCE   Homo sapiens.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 7686)
  AUTHORS   Holmskov,U., Mollenhauer,J., Madsen,J., Vitved,L., Gronlund,J.,
            Tornoe,I., Kliem,A., Reid,K.B., Poustka,A. and Skjodt,K.
  TITLE   Cloning of gp-340, a putative opsonin receptor for lung surfactant
            protein D
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (19), 10794-10799 (1999)
  MEDLINE   99415938
   PUBMED   10485905
REFERENCE   2 (bases 1 to 7686)
  AUTHORS   Holmskov,U., Mollenhauer,J., Madsen,J., Vitved,L., Groenlund,J.,
            Tornoe,I., Kliem,A., Reid,K.B.M., Poustka,A. and Skjoedt,K.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-JUN-1999) Department of Immunology & Microbiology,
            Institute of Medical Biology, University of Southern Denmark,
            Odense University, Winsloewparken 21.1, Odense 5000, Denmark
FEATURES   Location/Qualifiers
     source   1..7686
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q25.3-26.1"
     gene   1..7686
                     /gene="DMBT1"
                     /note="deleted in malignant brain tumor 1"
     CDS   107..7348
                     /gene="DMBT1"
                     /note="scavenger receptor cysteine-rich protein SRCR;
                     putative receptor for SP-D"
                     /codon_start=1
                     /product="gp-340 variant protein"
                     /protein_id="AAD49696.1"
                     /db_xref="GI:5733598"
                     /translation="MGISTVILEMCLLWGQVLSTGGWIPRTTDYASLIPSEVPLDQTV
                     AEGSPFPSESTLESTAAEGSPISLESTLESTVAEGSLIPSESTLESTVAEGSDSGLAL
                     RLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNAWFGQG
                     SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLRPESWPVR
                     ISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLG
                     CGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICS
                     APQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSW
                     DTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSH
                     NCQHSEDAGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVE
                     VLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGN
                     ESYLWSCPHNGWLSHNCGHSEDAGVICSGPESSLALRLVNGGDRCQGRVEVLYRGSWG
                     TVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCP
                     NNGWLSHNCGHHEDAGVICSAAQSRSTPRPDTLSTITLPPSTVGSESSLTLRLVNGSD
                     RCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLD
                     DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSVSQSRPTPSPDTWPTSHASTAGS
                     ESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGCGWATSAPGN
                     ARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAHSWSTPSP
                     DTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVV
                     CRQPGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYPWSCPHNGWLSHNCGHSEDA
                     GVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTV
                     CDDYWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHN
                     GWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQG
                     RVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRC
                     SGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSL
                     ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFG
                     QGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWP
                     TSRASTAGSESTLALRLVNGGDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGC
                     GWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSA
                     AQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDS
                     WDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLT
                     HNCGHHEDAGVICSATQINSTTTDWWHPTTTTTARPSSNCGGFLFYASGTFSSPSYPA
                     YYPNNAKCVWEIEVNSGYRINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICN
                     DTRQIFTSSYNRMTIHFRSDISFQNTGFLAWYNSFPSDATLRLVNLNSSYGLCAGRVE
                     IYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVECSGT
                     ESTLWQCRNRGWFSHNCNHREDAGVICSGNHLSTPAPFLNITRPNTDYSCGGFLSQPS
                     GDFSSPFYPGNYPNNAKCVWDIEVQNNYRVTVIFRDVQLEGGCNYDYIEVFDGPYRSS
                     PLIARVCDGARGSFTSSSNFMSIRFISDHSITRRGFRAEYYSSPSNDSTNLLCLPNHM
                     QASVSRSYLQSLGFSASDLVISTWNGYYECRPQITPNLVIFTIPYSGCGTFKQADNDT
                     IDYSNFLTAAVSGGIIKRRTDLRIHVSCRMLQNTWVDTMYIANDTIHVANNTIQVEEV
                     QYGNFDVNISFYTSSSFLYPVTSRPYYVDLNQDLYVQAEILHSDAVLTLFVDTCVASP
                     YSNDFTSLTYDLIRSGCVRDDTYGPYSSPSLRIARFRFRAFHFLNRFPSVYLRCKMVV
                     CRAYDPSSRCYRGCVLRSKRDVGSYQEKVDVVLGPIQLQTPPRREEEPR"
BASE COUNT   1614 a   2141 c   2168 g   1763 t
ORIGIN

Alignment Scores:
Pred. No.:   8.66e-185   Length: 7686
Score: 2980.00   Matches: 618
Percent Similarity: 51.62%   Conservative: 177
Best Local Similarity: 40.13%   Mismatches: 505
Query Match: 36.62%   Indels: 240

```
DB:                    9              Gaps:        23
US-09-759-130B-381 (1-1453) x AF159456 (1-7686)

Qy     44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
Db    794 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGGTC 853

Qy     64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThr 83
Db    854 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 913

Qy     84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet---------PheArgPhe 100
Db    914 GTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTT 973

Qy    101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
Db    974 GGCCAGGGC---------TCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACAC 1024

Qy    121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
Db   1025 GAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGT 1084

Qy    141 GluAspValGlyValAsnCysTyrGly--------------------------------- 149
Db   1085 GAAGACGCTGGTGTCATCTGCTCAGCTCCCCAGTCCCGGCCGACACCCAGCCCAGATACT 1144

Qy    150 ---------------------------------GluAlaAsnLeuGlyLeuArgLeuVal 158
Db   1145 TGGCCGACCTCACATGCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTG 1204

Qy    159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
Db   1205 AATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACC 1264

Qy    179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
Db   1265 GTGTGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGT 1324

Qy    199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
Db   1325 GGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTC 1384

Qy    219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
Db   1385 CTGGATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGC 1444

Qy    239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyr--------- 255
Db   1445 TGGCTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCCCAC 1504

Qy    255 ------------------------------------------------------------ 255
Db   1505 TCCTGGTCGACGCCCAGTCCAGACACGTTGCCGACCATCACCTTACCTGCATCGACAGTA 1564

Qy    256 ---AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArg 274
Db   1565 GGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGA 1624

Qy    275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAla 294
Db   1625 GTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGACACCAAT 1684

Qy    295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314
Db   1685 GATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTTGGCCCCAGGAAAT 1744

Qy    315 ProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334
Db   1745 GCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGACGTGCGCTGCTCAGGGAAT 1804

Qy    335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354
Db   1805 GAGTCCTACTTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGTGGCCATAGT 1864

Qy    355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373
Db   1865 GAAGACGCTGGTGTCATCTGCTCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 1924

Qy    374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIle 393
Db   1925 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCTTGGGGCACCGTG 1984

Qy    394 CysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysPro 413
Db   1985 TGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGC 2044

Qy    414 PheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIle 433
Db   2045 TGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTG 2104

Qy    434 AsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLys 453
Db   2105 GATGATGTGCGCTGCTCAGGACATGAGTCCTACCTGTGGAGCTGCCCCAACAATGGCTGG 2164

Qy    454 AlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSer------------ 469
Db   2165 CTCTCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCCCAGTCC 2224

Qy    469 ------------------------------------------------------------ 469
Db   2225 CGGTCGACGCCCAGGCCAGACACGTTGTCGACCATCACGTTACCTCCATCGACAGTAGGA 2284

Qy    470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489
Db   2285 TCTGAATCCAGTTTGACCCTGAGGCTGGTGAATGGAAGTGACAGGTGTCAGGGCCGAGTA 2344

Qy    490 GluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509
Db   2345 GAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGATACCAATGAT 2404

Qy    510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529
Db   2405 GCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCC 2464

Qy    530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGlu 549
Db   2465 CGGTTTGGCCAGGGCTCAGGACCCATTGTTCTGGATGATGTGCGCTGCTCAGGACACGAG 2524

Qy    550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569
Db   2525 TCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAA 2584

Qy    570 AspValIleValThrCysSer-----------------------GlyAspAlaThrTrp 581
Db   2585 GATGCTGGTGTCATCTGCTCAGTTTCCCAGTCCCGGCCGACACCCAGTCCAGATACTTGG 2644

Qy    582 -----------------------------------------GlyLeuArgLeuValGly 587
Db   2645 CCGACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 2704

Qy    588 GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal 607
Db   2705 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTG 2764

Qy    608 CysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAspCysPro 627
Db   2765 TGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGTGGC 2824

Qy    628 SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu 647
Db   2825 TGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTG 2884

Qy    648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667
```

```
Db   2885 GATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGG 2944

Qy    668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla------ 685

Db   2945 CTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCCCACTCC 3004

Qy    685 ------------------------------------------------------------ 685

Db   3005 TGGTCGACGCCCAGTCCAGACACATTGCCGACCATCACCTTGCCTGCATCGACAGTAGGA 3064

Qy    686 ------SerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysVal 703

Db   3065 TCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTG 3124

Qy    704 GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle 723

Db   3125 GAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGCGATGACAGCTGGGACACCAATGAT 3184

Qy    724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743

Db   3185 GCCAATGTCGTCTGCAGGCAACCGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCC 3244

Qy    744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763

Db   3245 CGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAG 3304

Qy    764 AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783

Db   3305 TCTTACCCGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATAGTGAA 3364

Qy    784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro--------------------- 795

Db   3365 GACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCGGCCAACACCTAGTCCAGACACTTGG 3424

Qy    796 ------------------------------------------------ArgLeuValGly 799

Db   3425 CCAACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 3484

Qy    800 AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal 819

Db   3485 GGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTG 3544

Qy    820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839

Db   3545 TGTGATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGC 3604

Qy    840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859

Db   3605 TGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTG 3664

Qy    860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879

Db   3665 GATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGG 3724

Qy    880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899

Db   3725 CTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCC 3784

Qy    899 ------------------------------------------------------------ 899

Db   3785 CAGCCGACACCCAGCCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCTGAA 3844

Qy    900 -----------ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915

Db   3845 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTC 3904

Qy    916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935

Db   3905 CTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 3964

Qy    936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle 955

Db   3965 GTGGTTTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTT 4024

Qy    956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975

Db   4025 GGCCAGGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTAT 4084

Qy    976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrVal 995

Db   4085 CTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCT 4144

Qy    996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015

Db   4145 GGTGTCATCTGCTCAGCTTCCCAGTCCCAGCCGACA---CCCAGCCCAGACACTTGGCCA 4201

Qy   1016 AspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg 1035

Db   4202 ACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTG-----------------GCC 4243

Qy   1036 LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055

Db   4244 CTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGC 4303

Qy   1056 PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075

Db   4304 TCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGG 4363

Qy   1076 LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095

Db   4364 CAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCA 4423

Qy   1096 GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115

Db   4424 GGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGC 4483

Qy   1116 ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCys 1135

Db   4484 CCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGC 4543

Qy   1135 ------------------------------------------------------------ 1135

Db   4544 TCAGCTTCCCAGTCCCAGCCGACACCCAGCCCAGACACTTGGCCAACCTCTCGTGCATCA 4603

Qy   1136 ---------SerGluPheThr---AlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151

Db   4604 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGGTGAAT------GGAGGTGACAGG 4657

Qy   1152 CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn 1171

Db   4658 TGTCGAGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGTGATGACTAC 4717

Qy   1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191

Db   4718 TGGGACACCAATGATGCCAACGTGGTCTGCAGGCAGCTGGGCTGT---------GGCTGG 4768

Qy   1192 ValSerLeuAlaProLeuSerLys---------ThrGlySerGlyPheMetTrpValAsp 1208

Db   4769 GCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGAT 4828

Qy   1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227

Db   4829 GATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTC 4888

Qy   1228 ------------------------------------------------GluArgArgIle 1231

Db   4889 TCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCTCAGTCCCAG 4948

Qy   1232 SerSerProAlaGluGluThrTrpIleThr-------------------------Cys 1242

Db   4949 TCAACGCCCAGGCCAGATACTTGGCTGACCACCAACTTACCGGCATTGACAGTAGGATCT 5008

Qy   1243 GluAspArgIleArgValArg---------GlyGlyAspThrGluCysSerGlyArgVal 1259

Db   5009 GAATCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTGTCGAGGCCGAGTG 5065
```

```
Qy    1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279
           |||:::   :::   ||||||||||||||||||||||||||||||||||||      :::
Db    5066 GAGGTCCTGTATCGAGGCTCCTGGGGAACCGTGTGTGATGACAGCTGGGACACCAATGAT 5125

Qy    1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAla 1299
           |||    |||||||||:::||||||||||||||    |||::::::|||      :::|||
Db    5126 GCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCGGCCCCAGGAAATGCC 5185

Qy    1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
              ||||||||||||:::|||   |||   |||||||||:::||||||   |||||||||
Db    5186 CGGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGGAATGAG 5245

Qy    1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
           |||:::||||||   |||         |||   |||         :::|||||||||   |||
Db    5246 TCCTACCTGTGGAGCTGCCCCCACAAAGGCTGGCTCACCCACAACTGTGGCCATCACGAA 5305

Qy    1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
           ||||||||||||   ||||||            :::   |||               |||
Db    5306 GACGCTGGTGTCATCTGCTCAGCCACCCAAATAAATTCTACTACGACAGATTGGTGGCAT 5365

RESULT 15
AC131207/c
LOCUS       AC131207              202252 bp    DNA     linear   HTG 18-AUG-2002
DEFINITION  Homo sapiens clone RP11-326A16, *** SEQUENCING IN PROGRESS ***, 10
            unordered pieces.
ACCESSION   AC131207
VERSION     AC131207.1  GI:22296912
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 202252)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202252)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HELB
                Center clone name: RP11-326A16
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 194619 bases at least Q40
                Consensus quality: 196802 bases at least Q30
                Consensus quality: 198233 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2422: contig of 2422 bp in length
            *     2423     2522: gap of unknown length
            *     2523     7722: contig of 5200 bp in length
            *     7723     7822: gap of unknown length
            *     7823    13397: contig of 5575 bp in length
            *    13398    13497: gap of unknown length
            *    13498    22395: contig of 8898 bp in length
            *    22396    22495: gap of unknown length
            *    22496    39888: contig of 17393 bp in length
            *    39889    39988: gap of unknown length
            *    39989    58697: contig of 18709 bp in length
            *    58698    58797: gap of unknown length
            *    58798    81609: contig of 22812 bp in length
            *    81610    81709: gap of unknown length
            *    81710   106273: contig of 24564 bp in length
            *   106274   106373: gap of unknown length
            *   106374   143244: contig of 36871 bp in length
            *   143245   143344: gap of unknown length
            *   143345   202252: contig of 58908 bp in length.
FEATURES             Location/Qualifiers
     source          1..202252
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-326A16"
BASE COUNT     60981 a  37684 c  37424 g  62900 t   3263 others
ORIGIN

Alignment Scores:
Pred. No.:               3.25e-170      Length:          202252
Score:                   2783.00        Matches:         1103
Percent Similarity:      18.39%         Conservative:    77
Best Local Similarity:   17.19%         Mismatches:      199
Query Match:             34.20%         Indels:          5044
DB:                      2              Gaps:            61

US-09-759-130B-381 (1-1453) x AC131207 (1-202252)

Qy      28 CysIleLeuLeuLeu---------AsnSerCysPheLeuIleSerSerPheAsnGlyThr 44
            |||:::|||:::              |||   ||||||   :::|||      :::||||||
```

```
Db   83194 TGCCTTCTAATCTTTGCCATCTCTAATGTATGTTTCTGTCTTTCTTGCA--GATGGAACA 83137

Qy      45 AspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLys 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83136 GATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAA 83077

Qy      65 PheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrVal 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83076 TTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTC 83017

Qy      85 ValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPheGlyGlnAlaVal 104
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83016 GTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTG 82957

Qy     105 ThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeu 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82956 ACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTC 82897

Qy     125 TrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGly 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82896 TGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGT 82837

Qy     145 ValAsnCysTyr------------------------------------------------ 148
           ||||||||||||
Db   82836 GTGAACTGTTATGGTAAGAGTCTAACTAGAGTAGGAAGGCAAGGTAAGACCTACCCTCAG 82777

Qy     148 ------------------------------------------------------------ 148

Db   82776 GAATCATCTGCTGGAAGGTAGGGGTTTCTATTATGACTTTAATAGTGTCCCTCTGGGAGG 82717

Qy     148 ------------------------------------------------------------ 148

Db   82716 TGTAGGAAATATTTTATATATTCTACCCTTCCCTTATTTCCAAAGTAATGTCTAATTGCT 82657

Qy     148 ------------------------------------------------------------ 148

Db   82656 TTTCTTTTTATAACTATAACTGAGTTTAGAAAGAAATGGAGAGAGAAAGAGAGAGACAGA 82597

Qy     148 ------------------------------------------------------------ 148

Db   82596 GAGAGAATGCATTACAGTTACCAATGTAATATCTTTATCATAACTATAATTGAGTTTAGA 82537

Qy     148 ------------------------------------------------------------ 148

Db   82536 AAGAAATGAAGAAGTTTGGGAGCAGGGTTAGCTTTGATTTAAGTGCCTTTTTTCATGTAC 82477

Qy     148 ------------------------------------------------------------ 148

Db   82476 CAGGTACACTAACAGAACAAATATTTTCTGAAAAATGAGGTATAAAATTAAAGCAAAATG 82417

Qy     148 ------------------------------------------------------------ 148

Db   82416 ATAAAGAATATCATGCTAATGATTGATTTTTTCACCTAGGATTTTTTTTGTTTGTGTATT 82357

Qy     148 ------------------------------------------------------------ 148

Db   82356 TGTGTTTAGAAAAAAAGCCTCACATACCCCTCCTATAAGTATTTTCAGTTATTTTTTTTTA 82297

Qy     148 ------------------------------------------------------------ 148

Db   82296 ATTTTTGGTAACTTCATTCATAACAGGGCTGTCTTCTACGTATCCTTTTGATTTTATTTC 82237

Qy     149 ---------------------------------------------------GlyGluAlaAs 152
                                                              |||||||||||
Db   82236 AGGCTCAAAGTTGAATCAGTTTTTCATATGTCTGCTTGTCTTTCTCGTAGGTGAAGCCAA 82177

Qy     152 nLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPh 172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82176 TCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATT 82117

Qy     172 eGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValVa 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82116 CCAAGAAAGGTGGGGAACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGT 82057

Qy     192 lCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAl 212
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82056 GTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGC 81997

Qy     212 aValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTr 232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81996 TGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTG 81937

Qy     232 pAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLe 252
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81936 GAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATT 81877

Qy     252 uThrCysTyrAspSerSer---------AspLeuGluLeuArgLeuValGlyGlyThrAs 269
           |||||||||       :::          :::|||  |||                 ||
Db   81876 AACTTGTTATGGTAAGAACCCAAGACATCATGTAGAAGAAAGGAACAATTTGTTTAGAAA 81817

Qy     269 nArgCysMetGlyArg-------------------------------------------- 274
           |         |||:::
Db   81816 CAGTATGTATGGAAAGAAAAAGGTCAATCATTCTCTGGAGAAGAATTTTCTCTGGTTATA 81757

Qy     275 ----------ValGluLeuLysIleGlnGlyArgTrpGlyThrVal-------------- 286
                     |||   ||||||:::             ::::::
Db   81756 ACACATTTTTGTTACCCTGAAATTGCTTGATGAACTACTGTCCATCGNNNNNNNNNNNNN 81697

Qy     286 ------------------------------------------------------------ 286

Db   81696 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 81637

Qy     287 ------------------------------CysHisHisLysTrpAsnAsnAlaAlaAl 296
                                         |||   |||                :::
Db   81636 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNATGATGTTTCCACTTCCAGATATAGATTTCCTT 81577

Qy     296 aAspValValCys----------------------------------------------- 300
                |||   |||
Db   81576 AAGTGTTTCTTGTATGACCAATCTAGTGGTGATGAATTTCCTCATTTTTTGCTTATCTGG 81517

Qy     301 -LysGlnLeuGlyCysGlyThrAlaLeu-------------------------------- 309
             |||   |||         :::  |||
Db   81516 GAAAGACCTTATATCTCCTTCATTTTTGAAGGGTAACTTTGCTGGGTATTATATTTTAGG 81457

Qy     310 ----HisPhe-------------------------------------------------- 311
               ||||||
Db   81456 CTGGCATTTTTTTCTTTTAGCACTTTGAATACACCATACCTTTCTTTCCTGGCCTGTAAG 81397

Qy     312 -------AlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyVa 329
                  |||   ||||||                 ||||||:::   :::|||     ::
Db   81396 ATTTCTGCTGGGAAATCCCCAT---------TCATCTGATCTGGGTTTCCTTGCACACAT 81346

Qy     329 lSer-------------------------------------------------------- 330
           :|||
Db   81345 GAGTGGCCATCTTTTTCTTGTATTTAAAATTATTTTTTTATCTTGGGCTTTTCACAGTTT 81286

Qy     331 ----------CysSerGlyAsnGluSerPheLeuTrpAsp-------------------- 340
                     |||   |||   |||      ||||||:::
Db   81285 GAGTATAATGTGCCTTGGA---GAATACATTTTGTGGGAGGAGGTGTATCAATTTGGAGA 81229

Qy     340 ------------------------------------------------------------ 340

Db   81228 TCTTTCAGCTTCTTGTATCTGGATATCTAAGTATCTTTTAAAACCTGGAAAGTTTCAGCT 81169

Qy     340 ------------------------------------------------------------ 340

Db   81168 ATTATTTTATTAAATAAGTTTTCTATGCCATTGTCCATCTCTTCTCCTTCTGGAACACCC 81109

Qy     340 ------------------------------------------------------------ 340

Db   81108 AGAATCCAAGTATTTGGTCACCTTATGGTGTCCCATGTGTTGCAAAGGTTTTTTAATTGA 81049

Qy     340 ------------------------------------------------------------ 340

Db   81048 AAAAAAATTTCTTTCTTTTTCTTTTACTAGGTTATTTCAAAGATTTATCTTCAAGTTCTG 80989
```

```
Qy      340 ------------------------------------------------------------ 340

Db    80988 CAGTTCTTTCTTCTGCTTTATTTGATTTTTTTATGTTTTTCATTGAATTATTTCATTCC 80929

Qy      340 ------------------------------------------------------------ 340

Db    80928 ATGATTTCTGTTTGTTTCTTTTTTATGATATTTCTTTGTTGAATTTCTCATTAACATCCT 80869

Qy      340 ------------------------------------------------------------ 340

Db    80868 AAATTGTTTTTCTAATTTATTTGTATTATTTATCTGTGTTTTTTTGTGTGTTTCACTGAA 80809

Qy      340 ------------------------------------------------------------ 340

Db    80808 ATTCTTTAATAACATTATTTAAAATTATTTTCAGGCAGTTCATCCATTTTCTTTTTGCTG 80749

Qy      340 ------------------------------------------------------------ 340

Db    80748 AAATGTGGTGCTGGAGAAATTATTGTGTTCCTTTGAAGGTGTCATGTTTCCTTGCCTTTT 80689

Qy      340 ------------------------------------------------------------ 340

Db    80688 TATGTTTCTTGTGTCCTTACATTGATATCTGTGCATCTGGTATAAAATCGTTGCTTTCAT 80629

Qy      341 -------------------------------------CysArgHisSerGlyThrValAs 348
                                                 ||||||:::|||    :::|||
Db    80628 ATTTTGGGGTTGGCTTTCATAGAGGAAAACTTTTTGTTGTAGATATAGCTATAGTGTTGG 80569

Qy      348 nPhe-------------------------------------------------------- 349
             :::
Db    80568 TTGGGTAGGGCAGTTTATCTTTGATCCTGGGTGCATGCTATAGTGTAATCTCCATGATTT 80509

Qy      350 ---------------------AspCysLeu------------------------------ 352
                                 :::||||||
Db    80508 CTTCAGTGATAATCAGTGTGAAGAGTGTCTGGGATTTCCTCTTTGGCTTAGACTGTGGCT 80449

Qy      353 ---------------------------------------HisGlnAsnAspValSerVa 359
                                                    ||||||        |||::
Db    80448 ATTAGTTAAGGCTGTTGTGAAGCTTTGCTAGGAATAAGGACACCAGACAAGCCAGTCCTT 80389

Qy      359 lIleCysSerAspGlyAlaAspLeuGluLeuArgLeu---AlaAspGlySer-------- 375
             :     |||   |||            |||   :::       |||:::
Db    80388 GAGCCTCAGTGGTGGTGGTAGCAGCAGGCTGACCATGACTGTCCATGGACCCTAGGTGG 80329

Qy      375 ------------------------------------------------------------ 375

Db    80328 CATATGCAAGCACCAATGTGAGTTGATCTAGGTGGGTCAATTCTTGGGCTTCCAGGTGGC 80269

Qy      376 -------AsnAsnCysSerGlyArgValGluVal-------------------------- 384
                   :::      ||||||   :::|||:::
Db    80268 TTGCTTATCTACCAGCAGTGGCAGCATTGAGTTAGGCAGGTGGGAAAGTTCTCAGGCCCC 80209

Qy      385 -ArgIleHisGluGlnTrp----------------------------------------- 390
             |||   |||:::   |||
Db    80208 TAGGTAACATGAC---TGGTGTCACCAATAGCAGTAGTGGTGTCAGGCCAACTGTCAGGT 80152

Qy      391 ---------TrpThrIle-CysAspGlnAsnTrpLysAsnGluGln------------- 402
                     :::|||:::  |||   |||   |||
Db    80151 TCCCAGCCGGTATACACTGGTGCTAGCAGTGCTGGGCAGGCAATTCTCCAGGCCCCCAGG 80092

Qy      403 ----------------------------------------AlaLeuValValCysLysG 409
                                                    |||:::   ||||||    |
Db    80091 TGGCACATGTAGGTGGGTGCTGGAGGTAGTGGTGGCACCCAGCAGTGCAGGTCTGTCCTC 80032

Qy      409 lnLeuGlyCysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaA 429
            ||   |||                   |||                       :::
Db    80031 AA---GGCTACTGATGTTGCACACAGGGACAGACTATGGTGAGTGGAATAGGGCAATTTC 79975

Qy      429 rgAspIleTrpIleAsnSerIleSer-CysThrGlyAsnGluSer--------------- 443
              |||:::|||:::        :::          |||        :::
Db    79974 CAGACCTTTGGGTGGCATGCTTAGGTGGTAGCAGGTGGGGCAAACCTCCATGTAGGTCCT 79915

Qy      444 ------AlaLeuTrpAspCysThrTyr---------AspGlyLysAlaLysArgThr--- 457

                    |||   |||           :::         :::|||      |||   |||
Db    79914 TGTATGGCATCCTGGCAAGCCATTTTCAGGCTCACTGAAGGCTCATGAAAGTGCACAGTG 79855

Qy      458 ------CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAsp 475
                        |||   |||          :::      ||||||       :::
Db    79854 TCCCTGCTGCTGCGAGGAAGTGGGGTTGCTATCAGTGATAGTGACCCCAGACAGGTGGCT 79795

Qy      476 LeuArgLeu---------------------------ValGlyAlaHisSerProCysTyr 486
            |||||||||                          :::|||   ||||||
Db    79794 CTCAGGCTCTGGGGAGATGCTTCAGCTCTCTTTATCATGGGGGAACACTCC--------- 79744

Qy      487 GlyArgLeuGluValLysTyrGln-------------------------GlyGluTrp 497
                      |||   |||                                  |||
Db    79743 ---------TCTGTGAACTACCCATTTCCCAGTACATCAGACACTGTGTGGAATAGATGG 79693

Qy      498 ---GlyThrValCysHis------------------------------------------ 502
               ||||||      |||
Db    79692 CTGGGGACCTAGACACACTGCCAGTCCAGCCAGCATGGCACCACTGCAGCCCTCCAGATG 79633

Qy      503 -----------AspArgTrpSerThrArgAsnAlaAlaVal------------------- 512
                       :::|||   :::|||:::      :::
Db    79632 GATAAGGGAGGTTATCAGTGGGGCTCTAGACATGTAGAGATGTAGGGCTTACTGGGGCCC 79573

Qy      513 ------ValCysLysGlnLeuGlyCysGly------------------------------ 520
                  |||      :::|||   |||
Db    79572 AGGACAGGATGCAATCTTGTGGGAGCTGGACTCTCAAAATGGTGCCATGCTGGAGTTACC 79513

Qy      520 ------------------------------------------------------------ 520

Db    79512 TGGGTCTTGGGGTTGGGGATTGGGATGTAGGACCCAGCACAAACTCCCCCTCTGGAACAA 79453

Qy      521 ------------------------------------------LysProMetHisValPhe 526
                                                      :::|||
Db    79452 CGCTATTGTGCAAACTCTAGGCAGCTCCCTGTACTAGTCTCAGAGCCCACAGAAGCTGAG 79393

Qy      527 GlyMetThrTyrPheLysGluAlaSerGlyProIleTrp--------------------- 539
            |||:::::::::      :::         ||||||   |||
Db    79392 GGGCTCTCCCATGGCTAGGATTGCAAGGGCCCACAGTGGGAATGTGGACTGCTAGGAATC 79333

Qy      539 ------------------------------------------------------------ 539

Db    79332 TCTCATTTATTCTTTTTCTACAGTGGGGATCCATACCAGGCTTCCAGTCAATCCTAGCCA 79273

Qy      539 ------------------------------------------------------------ 539

Db    79272 GACAATTGCTTCACTTCTTTCTCTCTCTGTGCCTCAGAAGGATCTGTTCCCTATCACTTC 79213

Qy      540 ------------------------LeuAspAspValSerCys------------------ 545
                                    ||||||   :::   |||
Db    79212 CCTGCTGGATTCCAGTGTTCTATCTTAGACACTATTCGTTGTGAAATTATCTACTTGCTG 79153

Qy      545 ------------------------------------------------------------ 545

Db    79152 TTTTGATCGTTTTTGGTGGTGAGTGTTGGGTGCCTCTAGTAGCCGTCTTGAAGCCTCCCC 79093

Qy      545 ------------------------------------------------------------ 545

Db    79092 CTTGAGCTATTCTCTTTTAAGGGAAAGTTCTTCTCTCCATAATCATGAATATCAATGTGT 79033

Qy      545 ------------------------------------------------------------ 545

Db    79032 CTTTCAATAAGAGAGTTGAGGATTCCTGAACACGGGACCCAACTCCAATCCATAGAAACT 78973

Qy      546 -------------------------------------------------------IleGl 547
                                                                   |||||
Db    78972 ACCCACACCTTTCTTCTCTGTCCTTGTATCTGTATTGTAAACTACAGACTTTCTTATTGG 78913

Qy      547 yAsnGlu----------------------------------------------------- 549
            ||||
Db    78912 TAATAATTGTTTTACATATTGAGAATTATATATATATCTCTATGTGTGTATATATATATT 78853

Qy      549 ------------------------------------------------------------ 549
```

```
Db  78852 TAAAATAATTATATATATAAAACAATTATATATAAACAATTCTATATAAAACCATTATAT 78793

Qy    549 ------------------------------------------------------------ 549

Db  78792 ATAAATATATAATTATGTATATAAAACAATTACATATAAATATATATTTTTTTAATTCAG 78733

Qy    549 ------------------------------------------------------------ 549

Db  78732 AAAAACCTAAAAGTCAATATGATGTCACCTTTTTATGCTGAATCACCTACAGCAACTGTT 78673

Qy    549 ------------------------------------------------------------ 549

Db  78672 TTTCCACCTAAGACTCTCTCTCAGAGTGGAATATGTTGCTACAAAGCAGTGAAATATGTT 78613

Qy    550 ----------------SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHis-- 563
                          |||:::   ::::::   |||   |||        :::|||
Db  78612 TTTCCCTTATATTTAGAGTAGTTCATTTAATGAAGAAAGTTCATCTGTTTCTAGACACCT 78553

Qy    563 ------------------------------------------------------------ 563

Db  78552 TTTCAGGGAATCTCAATTTTGATAGAGATTCCTACAAGATATATATGTTTATTCTCCCCA 78493

Qy    563 ------------------------------------------------------------ 563

Db  78492 CCCTCCCCACCACCACACACAAAATTATAGCTCTCTAAACGTCAGTATCTAAAATAAACC 78433

Qy    564 -------------------------------------------AsnCysValHisArgGl 569
                                                         ::::::
Db  78432 CCATTTAAAAAAATTTCTTTTTACTTTTATGGTAAAGTTGCTAAGACTGATCTACCACCC 78373

Qy    569 uAspValIleValThrCysSerGlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySe 589
             :::|||:::                       |||:::|||:::          ::
Db  78372 TCACATGATTATTGTT------------------------TTGAAACTCCTGCCAATTAC 78337

Qy    589 rAsnArg------------------------Cys-------------------------- 592
          :||||||                        |||
Db  78336 AAACAGATCTTGACTCATGTTTCACAAACCCTGTCACTGAGCTCCAAAGACACAGAGTAA 78277

Qy    593 ----------------SerGlyArgLeuGluValTyrPheGlnGlyArgTrp-------- 604
                             ·|||::::::||||||         |||
Db  78276 CAAATAGGCTTAAAATGGAAATACATTACAACTCTATTTTTCTGCTGAATGGGAGTTTTA 78217

Qy    604 ------------------------------------------------------------ 604

Db  78216 TTTAGAGAATCATCTCATATCTACACAAGATTATATCAAAATAACCAATGTAGTAAAATT 78157

Qy    605 -------------------------------GlyThrValCysAspAspGlyTrp----- 612
                                         ||||||      |||
Db  78156 TCATTCAGCTGTTTTATAAGTGGCAGGAAGTGGGACTGA-AGTGATAGGAAAGGACAGAC 78098

Qy    612 ------------------------------------------------------------ 612

Db  78097 TTTCCAGGTCCTCTAGACCACGTTAATTATATTTTTCTTCTTCTCAAGGCCAATAGGAAG 78038

Qy    612 ------------------------------------------------------------ 612

Db  78037 TAACTTGGTCAGGTTTGATCTGGTCAAAGTCAGAGTTGATTTTCTGACATGTCAGATCTG 77978

Qy    612 ------------------------------------------------------------ 612

Db  77977 TCAACATTAGGCAGTTAATAAACCTATTAATATAATTTAGTATATAAAGTAGAAAAATAG 77918

Qy    613 ----------------------------------------------------AsnSerLy 615
                                                              |||   ||
Db  77917 TATGATATCACATAGACATTTAATAATGTGAATTTTTATTTTTAACTAAAACAATAGAAA 77858

Qy    615 sAlaAlaAlaValValCysSer------------------GlnLeuAspCys-------- 626
          |   |||      :::   |||                   |||   |||
Db  77857 AGAAGCTGATACCATTTATAGTTTTTAATTAGATAAAGAAGTATTACATTGTTATATAAT 77798

Qy    626 ------------------------------------------------------------ 626

Db  77797 TAGATATATTCTCAAATTAGTAGGCAATATTATTTGTGTTTCACAGACTATGCCACAGCA 77738

Qy    626 ------------------------------------------------------------ 626

Db  77737 CACTCAAAGCTACCAAACTGCCCTAAAAAGATAAAGTAACCTTTCTAGAATGAGATTCAT 77678

Qy    626 ------------------------------------------------------------ 626

Db  77677 TTCATTTTATGTCAATACTGCTAAAAATTTCCTTCAAAAGCTGACTCTTTCAATTAGGCT 77618

Qy    626 ------------------------------------------------------------ 626

Db  77617 CAGAATTAATTTCTGATTTAATTCTTCGTTTCCTCTTTTTAAAAAGTTTTTTATTTCCA 77558

Qy    626 ------------------------------------------------------------ 626

Db  77557 TAGGTTTTTAGGGAACAGATGGTAATAGTTACATGAGTAAGTTCTTTAGTTTTGATTTGT 77498

Qy    627 ----------------------ProSerSer----------------------------- 629
                                ||||||||| 
Db  77497 GAGATTTTGATGCACTCATCACCCAAGCAGTATACACTGAACCCAATTTGTAGTCTTTTC 77438

Qy    630 -------------------------------------------------IleIleGlyMe 633
                                                           ||||||   ||
Db  77437 TTCCTCAAACCCCTCGTACCCTTTCCCCTGAGTCCTCAAAGTCCATTGTATTATTCTTAT 77378

Qy    633 tGlyLeuGlyAsnAlaSerThrGly----------------------------------- 641
          |   |||         |||
Db  77377 GTCTTTGCATCGTCATAGCTTAGCCCCCCACTTATGAGTGAGAACATACAATATTTGGTT 77318

Qy    641 ------------------------------------------------------------ 641

Db  77317 TTCCATTCCTGAGTTACTTCACTTAGAATAATAGTCTTCAATTCCATCCATGTTGCTGTG 77258

Qy    642 -------------TyrGlyLysIleTrpLeuAspAspVal---SerCys----------- 652
                       :::      :::||||||      :::   ||||||
Db  77257 AATGCCATTAATTCATTCCTCTTTATGGCTGAGTAGTATTCCATCATGTGTATATATATA 77198

Qy    652 ------------------------------------------------------------ 652

Db  77197 CACACATACATGTGTGTATACGCATATACACACATACATGTATGTATACGCGTATACACA 77138

Qy    652 ------------------------------------------------------------ 652

Db  77137 CATACATGTATGTATACGTGTATACATACATATATGTATGTATATGTGTATACACACACA 77078

Qy    653 ----------------------------------------------AspGly-------- 654
                                                        ||||||
Db  77077 CACACACACACACACACACCACAGTTTCTTTATCCACTTGTTGATTGATGGGCATTTGGG 77018

Qy    654 ------------------------------------------------------------ 654

Db  77017 CTGGTTCCATATTTTTGCAATTGCGATTTATGCTGCTATAAACATGCGTGTGCAAGTATC 76958

Qy    655 ----AspGluSerAspLeuTrpSer----------------------------------- 661
                :::::|||||||:::|||
Db  76957 TTTTTCAGATAATGACTTATTTTCCTCTGGGTAGATACCTAGTAGTGGGATTGCTGGATC 76898

Qy    661 ------------------------------------------------------------ 661

Db  76897 AAATGGTAGCTCTGCTTTTAGTTCTTTAAAGAATCTCCACTGTTTTTCATAGCGGTTGTA 76838

Qy    662 ----------------------CysArgAsn----------------------------- 664
                                ||||||:::
Db  76837 CTAGTTACATTCCCACCAGCAGTGTAGAAGTGTTCCCTTTCCACTGCATCCATGAAAACT 76778

Qy    665 ----------------------------------------------SerGly-------- 666
                                                        ||||||
Db  76777 CTATTCTTTTTTGATTTTTTGGTTATGGCCATTCTTGCAGGAGTAAAGTGGCATTGCATT 76718

Qy    666 ------------------------------------------------------------ 666

Db  76717 GTGGTTTTGATTTGCATTTTCCTGATCATTAGTGATGATGAGCATTTTTCATTAAAAAAA 76658
```

```
Qy    666 ------------------------------------------------------------ 666
Db  76657 AAAAAACTTACCTACTCCTCTTCTCACATACAGTGAATTTTGTGTTTTCTTTTCCCGAGC 76598
Qy    667 ----------------------------TrpGlyAsnAsnAspCysSerHisSer----- 675
                                      |||   :::           |||   |||
Db  76597 TCCAATTTCTTTTTAGAATTACTTCATTTGGCATCACTAATAGGCAAGTTTCTCCATTTT 76538
Qy    675 ------------------------------------------------------------ 675
Db  76537 TCTGGATTTGGTAATATCACAGCCCTATGCAGTTTTTCCTTGACTTTTCATTAAATTATA 76478
Qy    675 ------------------------------------------------------------ 675
Db  76477 TCACCTCTTCCTAACATTCTCTTTAGTATCTAACATGTAATATGTAATTCTTCAAATATT 76418
Qy    676 ----GluAspValGly-------------------------------------------- 679
              :::::::::|||
Db  76417 CAGGCAAGAATTAGGGCCATTAAGTAACACTATTTTTTATCCTAAGTCTCTGATAAGATT 76358
Qy    679 ------------------------------------------------------------ 679
Db  76357 TTGTAAGTAAAACTATATCCTTGCTTTGTAAGCCAGAAATGTTCATGGATTGAAACATAC 76298
Qy    679 ------------------------------------------------------------ 679
Db  76297 CTTGGTGTACTTTTGTTATTTTATTTTTTATCCAATAAAAAGGAAAATTAAGTTTGAAGG 76238
Qy    679 ------------------------------------------------------------ 679
Db  76237 AAAGTAAAAAAATAAAATTAAACACTGGTGAAGCCGGGACCTAAACAATCTCTCAACTTTT 76178
Qy    679 ------------------------------------------------------------ 679
Db  76177 GGCCACTGGTCTATGTTAGTGTACAGGAATGGATAAGAATCTTTAATGTCTCACATAGAG 76118
Qy    680 ----------------------ValIleCysSerAspAlaSerAspMetGluLeuArgLe 692
                                   :::   ||||||||||||||||||||||||||||||
Db  76117 TGCTTCTCTAAAATCATTGTGGTATTTGGTTTCAGATGCATCGGATATGGAGCTGAGGCT 76058
Qy    692 uValGlyGlySerSerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGl 712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  76057 TGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGG 75998
Qy    712 yIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGl 732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  75997 AATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGA 75938
Qy    732 uCysGlySerAlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIl 752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  75937 ATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACAT 75878
Qy    752 eLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTr 772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  75877 CTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATG 75818
Qy    772 pGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSer----- 790
          |||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  75817 GGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTC-AGGTA 75759
Qy    790 ------------------------------------------------------------ 790
Db  75758 ATTTCTGGACAAAGATAAACACATTTATTTAATTATTTATTTTTGCACTTATTTAATACG 75699
Qy    790 ------------------------------------------------------------ 790
Db  75698 TGTTTATGAAGTGGCTATTTAACAGTATTAGGTCAAAAGTAAATGAACTAGATAAAAGTA 75639
Qy    790 ------------------------------------------------------------ 790
Db  75638 ATCCTACCCTCTATCTAGAAGACAGGTATATAATAAAAATATATAATTAAATAACTGTTT 75579
Qy    790 ------------------------------------------------------------ 790
```

```
Db  75578 ATTTAAAATTGCGTGAAGTAGAAGCACAAGATGAAATGAGAAGGCATAATTAGAAGATAT 75519
Qy    790 ------------------------------------------------------------ 790
Db  75518 TACCTATTTGCAGAGTCGGAGAATATTTCTGTGAAGTGATAGTTAAGGAAAGATGTGAAG 75459
Qy    790 ------------------------------------------------------------ 790
Db  75458 GACAGGTAAGATTTAATCATATGAAGAAGTGTTGGGAAGATTAATTTAAACAGTAGTATT 75399
Qy    790 ------------------------------------------------------------ 790
Db  75398 TTATTCGAGTAACACACCTGGGAATAAGCTTGTGTTATTCCAGGAACTAAAAGATGGTTG 75339
Qy    790 ------------------------------------------------------------ 790
Db  75338 AAAGAAGAGGTATTAGATCCAAGTGAAGAAGGGGAAAACTTATAATAAAGTCTTGATAT 75279
Qy    790 ------------------------------------------------------------ 790
Db  75278 GTAGGCAAGGGAAAGGTTATTTATGGAGTAGGAATTGTGAACTTTATTCTAAAGTGAATG 75219
Qy    790 ------------------------------------------------------------ 790
Db  75218 ATAAATCATTGAAGGTTGTAAGCATGATACAACTTGCATTTCAAAGACAGCACTCTGGCT 75159
Qy    790 ------------------------------------------------------------ 790
Db  75158 GCAGTGTGGGGTGTGTGTGTGTGTTTATGTGTGTGTCTGTGTGTGCATTGGAGAGAGGAG 75099
Qy    790 ------------------------------------------------------------ 790
Db  75098 GCTATTGGATGCCGGGAGTCATGCTGTTGAATAATTCTAGGGAAGATACGGAATTTTGAT 75039
Qy    790 ------------------------------------------------------------ 790
Db  75038 TAGGATGCTCACTGTGGAGCTGGTGAAAAGCTTCCAAATTTCAGAAATATTTAGTAGCTG 74979
Qy    790 ------------------------------------------------------------ 790
Db  74978 AAACCAGCAGAACTGAATACACTGAATGTGGAAATAAAGATAAAGGTTTCTAGACTAAGC 74919
Qy    790 ------------------------------------------------------------ 790
Db  74918 AGATGGGTGCACTGCAGTGTCATCATCTGGAAAGATAATCTTGGAAGAGGAGCAGTTTAG 74859
Qy    790 ------------------------------------------------------------ 790
Db  74858 AGAGGGTGGGATGATGAGTCCAATTTTGTATGTGTGAAATATGAGGTAATAGTAAAGCAG 74799
Qy    790 ------------------------------------------------------------ 790
Db  74798 TCATCAAATGAAGAGGGTTGGAGGAAATTGGATATGAGATTTAGATGATCAGAAGACTTG 74739
Qy    790 ------------------------------------------------------------ 790
Db  74738 TATTGAATGGAGATATAACTATTAGAGTCCTACATATGTAGACGGAAACTAAATCTAGGA 74679
Qy    790 ------------------------------------------------------------ 790
Db  74678 ATAAAATGTAAATTGAGAATTTTTAAAGCCTAGATTCAAGAAAGAGGAGCTCTTAGCCCT 74619
Qy    790 ------------------------------------------------------------ 790
Db  74618 GAGGAACAATAGTATTTAATAGTTTATTGACATAAAAGAAGTTGGAAAAGCATACTGAGG 74559
Qy    790 ------------------------------------------------------------ 790
Db  74558 AACAGTGGTGGCAGAATGTCACATCGCTAAAGCCAAGGGAAGATGTTTCAAAAAGATGGG 74499
Qy    790 ------------------------------------------------------------ 790
```

```
Db  74498 GACATCAATTCCAAAAAAGCCCTAAGAGGCCAATATAGATGAAACTGAACAATTTCCACT 74439
Qy    790 ------------------------------------------------------------ 790
Db  74438 AAGCTTAGCTACAGAGAAGTCATGAATAAATTTAATAAAAGTAGTTATGTCACATTAGAA 74379
Qy    790 ------------------------------------------------------------ 790
Db  74378 AAAGGAGATCATTTCAAGTTGAGAAGGAAATGGGAGGTAAAGATGTTGGGATAATAGGCA 74319
Qy    790 ------------------------------------------------------------ 790
Db  74318 TAGACTATTTCCAAAAAGCTGTATTTTAAGGGGAAGAAATAAATAAGGTGATGACTAAAA 74259
Qy    790 ------------------------------------------------------------ 790
Db  74258 AGGAATATAATCTTCAGGGGAGTTACTCAATTTTTTTATTGATAGGAGATTCTAGAGTA 74199
Qy    790 ------------------------------------------------------------ 790
Db  74198 GTGCTGTTCAAAAGAACATCACTTTCTGTGATGAAGTAAAATATATGCTATTAGCTGATG 74139
Qy    790 ------------------------------------------------------------ 790
Db  74138 TTAGAGGCTATAGGAAAAAAATAGAACTAGAAAATAGAATCCAGAGTGTTTATGGGACAGA 74079
Qy    790 ------------------------------------------------------------ 790
Db  74078 ATGAAATTCCAAATAAAATGACTGGGGTAGGCCTTACTGAAAAGGTGAAACATAGACAAA 74019
Qy    790 ------------------------------------------------------------ 790
Db  74018 GACTCGAAGGTCATGGGAGAGTCAGTTATATGCATATACGCAGAAGGAAAATACAAATAG 73959
Qy    790 ------------------------------------------------------------ 790
Db  73958 AGATACAGCCAGTTCAAAAGCCTTGGGGCAAGATTATGCCTGATATGTTTGAAGAAGAGA 73899
Qy    790 ------------------------------------------------------------ 790
Db  73898 AAGGAATCCAATGCAGCTGAAGAGCAGTAAGTAAAACATCAGGACCTGGCACCAAAAGGG 73839
Qy    790 ------------------------------------------------------------ 790
Db  73838 GTAAATGAGCAGGGGAAGGTCACTCAGGGACAAGCAGGTCTTCGTAGGAGTTTAGATTTT 73779
Qy    790 ------------------------------------------------------------ 790
Db  73778 TACTCTGAGTGAAATATGGAGTCATGGGAGGATTTGGGGCAGAATACTGATTTAACCTGA 73719
Qy    790 ------------------------------------------------------------ 790
Db  73718 TACATCATTTTAAATAACCTCTCTGGATTTTTATTTTTTTAAAAGATGAACTATAGAGGA 73659
Qy    790 ------------------------------------------------------------ 790
Db  73658 ATAATTATAGAAACAGGGAATCTGTAGGAATCTATTATAGTAACTCAGGCAAGTAATGAT 73599
Qy    790 ------------------------------------------------------------ 790
Db  73598 GGTGGCTCAATCAAGGCGCTAGGGTTGAAGGGGTGAGAAGTTTTGGGATTATGGGCATAC 73539
Qy    790 ------------------------------------------------------------ 790
Db  73538 TTTGAAAATACAGAAAACAAAACTTTCTGAAAGGCGAGATATAGTATGTAATGGAAGTCA 73479
Qy    790 ------------------------------------------------------------ 790
Db  73478 AGAATTATTTCAAAATATTTGAATGTCTGGAATGATGGAGTTCCATCAATAGAAATATGG 73419
Qy    790 ------------------------------------------------------------ 790
Db  73418 AAGTCTGCAGCCTTAGATAGATGTTTGAAAGGACTCCAAAGAATTCACATTGAGATATCT 73359

Qy    790 ------------------------------------------------------------ 790
Db  73358 ATAAGAAATCCAAATGGAAATGCTAAGTAGAAATTGGATAAATGAGTCAGGAGTTTGGAA 73299
Qy    790 ------------------------------------------------------------ 790
Db  73298 AAATGCCTGGGTTGGAGACAGATAACATACATTACTAATATAAGTAGCTTTGCTGCAAAG 73239
Qy    790 ------------------------------------------------------------ 790
Db  73238 AATAGAGTCAGAATAATGGGCAGTATTTTTTTGTTGTTTATAAATTACTCATTTAGATGG 73179
Qy    790 ------------------------------------------------------------ 790
Db  73178 ATTGCTGGGATCTAACCAGTAAGTAGAAGTCTTTCTATTCTTATATTACACCATCTTCTT 73119
Qy    790 ------------------------------------------------------------ 790
Db  73118 AAGATGACTACAAATCTTCACGTGTTGTTTAATGCTGAGGTTACTTAGTATTTTTTTAAT 73059
Qy    790 ------------------------------------------------------------ 790
Db  73058 GTAAACAAATTGTGCAATTCATTTCCTTCTGATGAGTTCTAAATCTAGGATTGGTGGGTC 72999
Qy    790 ------------------------------------------------------------ 790
Db  72998 ACCAAGACATAGAAAATAGGTATGTAATCCAGATCTGTGAATTTAAGTCTGTACTGCTAA 72939
Qy    790 ------------------------------------------------------------ 790
Db  72938 TCTTGTCCTCTGAGCCCCTACGAATGGGATGTTCTAATGATGAATGGAGAAAAATTGCTC 72879
Qy    790 ------------------------------------------------------------ 790
Db  72878 TGGGGTTAGATATGGAGACTGGAGACAGGGATAGATATTGGAAAACCCTAAAAACTAACC 72819
Qy    791 ------------------AlaHisArgGlnProArgLeuValGlyAlaAspMetProCy 804
                            ||||||||||||||||||||||||||||||||||||||||||
Db  72818 TTTGAAATATTATTTTATAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 72759
Qy    804 sSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPh 824
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72758 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 72699
Qy    824 eSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLe 844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72698 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTGTGGAGATGCCATATCTCT 72639
Qy    844 uSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCy 864
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72638 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 72579
Qy    864 sGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCy 884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72578 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 72519
Qy    884 sIleHisSerArgGluValGlyValValCysSerArg---------------------- 896
          ||||||||||||||||||||||||||||||||||||||
Db  72518 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCG-TGAGATTTTAAATAATCTTAAT 72460
Qy    896 ------------------------------------------------------------ 896
Db  72459 TGGCTGGGGCTAGAGATGGGGAGTTTGGGGCTGGAGCTGATAAGTCATGATGTCTCTTCT 72400
Qy    897 ---------------------TyrThrAspValArgLeuValAsnGlyLysSerGlnCy 909
                               ||||||||||||||||||||||||||||||||||||||
Db  72399 ATGGCCTTGTCTCCTCCTAGGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTG 72340
Qy    909 sAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTr 929
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72339 TGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTG 72280
```

```
Qy      929 pAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerTh 949
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    72279 GGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAAC 72220

Qy      949 rThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCy 969
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    72219 CACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTG 72160

Qy      969 sLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCy 989
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    72159 CTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTG 72100

Qy      989 sIleHisGlyAsnThrValSerValIleCysThr-------------------------- 1000
            ||||||||||||||||||||||||||||||||||
Db    72099 TATCCATGGAAATACTGTCTCTGTGATCTGCACAGGTAAGAGAATAGTGCTTATGGTTAA 72040

Qy     1000 ------------------------------------------------------------ 1000

Db    72039 CTACTGTCAATATAGCCCAGACATGGTAGAGGGAGTTGGGATTAAAGGAAAGATGAGTAA 71980

Qy     1000 ------------------------------------------------------------ 1000

Db    71979 GAAAATGGGGGTGCAGGTGGAGGGGGGATAGGCATGACAGTCAAAGAGTAACATGTAAGA 71920

Qy     1000 ------------------------------------------------------------ 1000

Db    71919 AATAATGATAACTAGGACCATTCCTAAGGGGAAGTTAACACCTCTGTGTTTGGGATTTAC 71860

Qy     1000 ------------------------------------------------------------ 1000

Db    71859 GAGTAGACATCATTAATAAAAATGATTTCAATAGACGACAGGGTTTGAAAAGTCCTTATC 71800

Qy     1000 ------------------------------------------------------------ 1000

Db    71799 ACAATTGCAGATTCCATGTTGTTACAGAACTCTTCAACCTTTTCCATGTATTTTGGTCTT 71740

Qy     1001 -------------------------------GlySerLeuThrGlnProLeuPheProCy 1010
                                           |||||||||||||||||||||||||||||
Db    71739 ACTTTTGCCCCATTTCCTGCACCCTCCACCAGGAAGCCTGACCCAGCCACTGTTTCCATG 71680

Qy     1010 sLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCy 1030
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71679 CCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTG 71620

Qy     1030 sLeu-------------------------------------------------------- 1031
            |
Db    71619 CTT-AGGTAAGCAGAATCAGTTTTAGGAAACACTGCAACTCTAAAGAATGTGACCTATTA 71561

Qy     1032 -------------------GluAspLysArgLeuArgLeuValAspGlyAspSerArgCy 1045
                               |||||||||||||||||||||||||||||||||||||||||
Db    71560 ACTGGACTTTCTCCTCTAAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTG 71501

Qy     1045 sAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTr 1065
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71500 TGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTG 71441

Qy     1065 pAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAl 1085
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71440 GGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGC 71381

Qy     1085 aThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCy 1105
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71380 CACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTG 71321

Qy     1105 sThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCy 1125
            |||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71320 CACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTG 71261

Qy     1125 sArgHisLysGluAspAlaGlyValIleCysSer-------------------------- 1136
            ||||||||||||||||||||||||||||||||||
Db    71260 CAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGGTCTGTGCTGCTACCTTTTCACCCT 71201

Qy     1136 ------------------------------------------------------------ 1136

Db    71200 CCTCTACACACACGTGCACACACTTGCACACACACAACACAATTGTAAAGTGTAGTGAAA 71141

Qy     1136 ------------------------------------------------------------ 1136

Db    71140 AGTGTATGGCCTAGGAATACAAACAAAAAATCCACTTGGTTTTAATCCTATTCTCTTATA 71081

Qy     1136 ------------------------------------------------------------ 1136

Db    71080 CTGATTTCTTAATACCCTTGTCATATTACATTTGTCTGGGTATCATTCATAACTGAGAAG 71021

Qy     1136 ------------------------------------------------------------ 1136

Db    71020 AATAAAAGATTGGTCTTACTGTCAGAGAATGTTAGACCCCAAGAAAGGGACCTTACATGA 70961

Qy     1136 ------------------------------------------------------------ 1136

Db    70960 CATCCACTCTGGCTCTGAGAACACACTGTCAACCGTGTCCAATCTGCCTTATTTTGTTGT 70901

Qy     1136 ------------------------------------------------------------ 1136

Db    70900 TTTGTTTTGTTTTATTTTAGGCTTTTAGTAGCCTGAAGCCATGGTTATTAGTTTCTGTCT 70841

Qy     1136 ------------------------------------------------------------ 1136

Db    70840 CTAGTGGTAAGAGGAAAGAGGGATGAGGAAGGGTCTTTACTAACCTGACCAGAAGCAGAA 70781

Qy     1136 ------------------------------------------------------------ 1136

Db    70780 ACTAAGAACCCATAACTGTATCCTCTCCCTTGGACACCCCTGCAAATGAGGTTAACCTCC 70721

Qy     1136 ------------------------------------------------------------ 1136

Db    70720 TGATTATTTTCTGAATTGAATTTATGTTCGAGAGTAGAATGTAGAAGGTTGCATTTTAAC 70661

Qy     1136 ------------------------------------------------------------ 1136

Db    70660 TTAAGTTGATACAAATTTAAAGCTCAAATACTATTCTTGGGAAACCCACACTCATCCTGA 70601

Qy     1136 ------------------------------------------------------------ 1136

Db    70600 AATCTTACAAGACCCTTAGTTTCCATCAGATTTATGCTGTTAATATGTTAGAGTGATTTA 70541

Qy     1136 ------------------------------------------------------------ 1136

Db    70540 TGAAATTGTATTGTGAGCCCTAATGTATTATTTCCATGGGAAAAATGGATTCCTTGGGGA 70481

Qy     1137 -------------------------------------------------------GluPh 1138
                                                                   |||||
Db    70480 GGGATTTTCCAACTCCAAAGGAAATATTTAATACTAAGTTTTGTATTTCTCTGTAGAATT 70421

Qy     1138 eThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluVa 1158
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70420 CACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT 70361

Qy     1158 lPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGl 1178
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70360 CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGG 70301

Qy     1178 yIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSe 1198
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70300 CATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATC 70241

Qy     1198 rLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSe 1218
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70240 TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTC 70181

Qy     1218 rIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluTh 1238
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70180 CATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGAC 70121

Qy     1238 rTrpIleThrCysGlu-------------------------------------------- 1243
            ||||||||||||||||
```

```
Db  70120 CTGGATCACATGTGAAGGTGGGTATCTTTCTAAACACCAGCTGTCACTGAAGCCCTGTGA 70061
Qy   1243 ------------------------------------------------------------ 1243
Db  70060 AATATAACTTACCATAAATAGCAAATCTTTCCATTCAGGAGCACATGGCAGACTCACTGA 70001
Qy   1243 ------------------------------------------------------------ 1243
Db  70000 AAAGTGCATGACAATAAATGCTGATTTATCGCATCTGTTCTATATTGAAATTTTATTATT 69941
Qy   1243 ------------------------------------------------------------ 1243
Db  69940 TTGGTTTAGTATTTTACAGGGTATTTAAAAATGTATATTGTCTCTTTTGATTCCTTTAAG 69881
Qy   1243 ------------------------------------------------------------ 1243
Db  69880 CTCCTTGTGAGAAGCTGGATGAAAGCACATAAAAGGATAACAACACATAAAAAATTACAG 69821
Qy   1243 ------------------------------------------------------------ 1243
Db  69820 TGTTTTAAAAAATGAAAAATATAATATGTATACTGTCTTAGAAATTCTAGAAAGCTCAGA 69761
Qy   1243 ------------------------------------------------------------ 1243
Db  69760 ACTTTTGAGGAAATGTGAATTAATTATGGAGATTAAGCAGCTCTATTTTTTCCTGCAGAA 69701
Qy   1243 ------------------------------------------------------------ 1243
Db  69700 CATTTTACTATGGGAAGTTAGTTGATTATTTATAAGACACTTGTAAATTGTCTTATTTGA 69641
Qy   1243 ------------------------------------------------------------ 1243
Db  69640 TTAAAATTTAATATGTTTATTGTAGAGATTTTAGAAATTGTACAATTATAAATAATAAAA 69581
Qy   1243 ------------------------------------------------------------ 1243
Db  69580 TGCATTCCACGTTTAATCTGATTATCAAAAATAACCATTACTAATATTCAGAAACATTAA 69521
Qy   1243 ------------------------------------------------------------ 1243
Db  69520 CTATTAGTCTTCCTTTTGTACAAAAACTTAAAGTTATGATTGTGTATATGAAATGTGTAT 69461
Qy   1243 ------------------------------------------------------------ 1243
Db  69460 ATGAAATATAGATACTATATATGTATATGTATATATAGTATGATATATATGCATGTAATG 69401
Qy   1243 ------------------------------------------------------------ 1243
Db  69400 TAGTGTTACACTATATTTAGGATATCTGTATCCTACTTTTATATGTAAAAGTATACTGTA 69341
Qy   1243 ------------------------------------------------------------ 1243
Db  69340 AATATATTCTAATTTCAATGAATAGCCTTCAAATGGCATACAATATTTTATCATGTGAAC 69281
Qy   1243 ------------------------------------------------------------ 1243
Db  69280 ATAAATAACCATGTGATTATTGGTATTAAGGCTATTTCTCAATGTCAAACTAGTATAAAT 69221
Qy   1243 ------------------------------------------------------------ 1243
Db  69220 AATGCCATTATAAGCATCCTTATTCATTTTCGGTGAATCTCTGATTATATCTGAAGGTTA 69161
Qy   1243 ------------------------------------------------------------ 1243
Db  69160 TATTTCTAGTCCCAAAATGAGTGGGTAAAATGAGTGAACATTTTTATGATTCTTTATATG 69101
Qy   1243 ------------------------------------------------------------ 1243
Db  69100 TACTAGTTAATTGCTCCAGCAGAAAGCCTATGTATTAGTTTTCATTCTTATCCCCAGTAT 69041
Qy   1243 ------------------------------------------------------------ 1243
Db  69040 ATGAGAGCATTGATTTTATTTCACCCTTCCCACCAATGTTACACACAGAGAGAAACATAT 68981

Qy   1243 ------------------------------------------------------------ 1243
Db  68980 ACACAGACATACACATACCTTCCCAAATTAAAATTTCACCTAAATCAACATAATTTGCAT 68921
Qy   1243 ------------------------------------------------------------ 1243
Db  68920 GGGTTTGGTTAGTAGTGGATATGGACATTATTTATGAATATTGTCTGTTTGTGGTTCATA 68861
Qy   1243 ------------------------------------------------------------ 1243
Db  68860 TTTTTAAGCATTTTGCTTCTTTTTACTATTGTAATTGTATCACAAAATAATATCATTCCC 68801
Qy   1243 ------------------------------------------------------------ 1243
Db  68800 CAAACAAAAAATTTGAAAAAAACCCAGAAATCATGGAAAAAAACACAGATATCACCTATA 68741
Qy   1243 ------------------------------------------------------------ 1243
Db  68740 ATACCATGATCAGTTGATGATATATTAGCTGATATTACACTCTTGAGTATCATATGTGTG 68681
Qy   1243 ------------------------------------------------------------ 1243
Db  68680 TATATATTTTATATATATTCATATATTTTATTGCAAAATGTTATGCTTGCATATTTGTGC 68621
Qy   1243 ------------------------------------------------------------ 1243
Db  68620 TCCCTGCTCTTTTGATTTAGCAATATTTTCTCAAGCTAAAGCATATATATTTTTTACCAT 68561
Qy   1243 ------------------------------------------------------------ 1243
Db  68560 ATGATATACTATAATCTTATGCACCAAATCCCTGTCATTGGCCATCTAGAAAATTCAAGT 68501
Qy   1243 ------------------------------------------------------------ 1243
Db  68500 TATTTTTCCATTTTAAGTAACATCACATTAAATATCCTTAAATATATAGAAAATTATTTT 68441
Qy   1243 ------------------------------------------------------------ 1243
Db  68440 AAGCATATTTCTCAATATTTTCTTCCCAAAATCTATTTTCCTATTTGATCAATTATATTA 68381
Qy   1243 ------------------------------------------------------------ 1243
Db  68380 CCTGAAATTTTGGATTTCAACTATCTTAAGTTTTAATTTATCAATATTTTATATTAAAAA 68321
Qy   1243 ------------------------------------------------------------ 1243
Db  68320 TTTGCATCTATATGTATATGAATGATACTATGCAGTAGATCTTATTTAGCGCTCTTTTTA 68261
Qy   1243 ------------------------------------------------------------ 1243
Db  68260 TATTTGTTGATATTAGCATTGGAGATTTGCTGAAATATAAACTAGGATGAATTCTGCCAG 68201
Qy   1243 ------------------------------------------------------------ 1243
Db  68200 CTTTTATGATCTTTAATAGTATATGGAATAAAATCATTGTGTTTCAATATTTGAAAGAGA 68141
Qy   1243 ------------------------------------------------------------ 1243
Db  68140 ATAACTATAGCTATTTTGGGTAGTAATTTGTTAATACTTAAAAAGAATGTCCAGGACTAT 68081
Qy   1243 ------------------------------------------------------------ 1243
Db  68080 TGGCTTATTCTCATTTTCTTGAGTCAATTTTGTGAAAGGATGAAAGAATGTATTAGTCAG 68021
Qy   1243 ------------------------------------------------------------ 1243
Db  68020 TTCTCATGCTGCTAATAAAGAAATACCAGAGACTGGGTAATTTATAAAGGAAGGAGGCTT 67961
Qy   1243 ------------------------------------------------------------ 1243
Db  67960 AATGGACTCACAGTTCCATATGGCTAATGAGGCCTCACAATCATGGCTGAAGGCAAAGGA 67901
```

```
Qy      1243 ------------------------------------------------------------ 1243

Db     67900 GGAGCAAAGTCACATCTTACATGGTGGCAGGCAAAAGAGCATGTGCAGGGGAACTCCCAT 67841

Qy      1243 ------------------------------------------------------------ 1243

Db     67840 GTATAAAACCATCAGATCTTGTGAGACTTATTCACTATCACAAGAACAGCACAGGAAAGA 67781

Qy      1243 ------------------------------------------------------------ 1243

Db     67780 CCCACCCCTATGATTCAATTACCTCCCACCAGGTCCCCCCACAACACGTGGGAATTATG 67721

Qy      1243 ------------------------------------------------------------ 1243

Db     67720 GGAGCTACAGTTTGAGATTTGGGTGGGGATACAGCCAAACCATAGCAAAGAAGTAGGATT 67661

Qy      1243 ------------------------------------------------------------ 1243

Db     67660 AAAAACCATACTCTGAATTTATTATTATATAGGCCATTCAATGATGCACTCTGTATTGGT 67601

Qy      1243 ------------------------------------------------------------ 1243

Db     67600 AGGAGGTAGAAGCTAGTCTAAAAGTATTTATTAATGAGCAGGAGATTTACAAAAATGACG 67541

Qy      1243 ------------------------------------------------------------ 1243

Db     67540 AGAAAAAATAAACTCAAGATGTTTGGCTATGAGAGAGAGAAAAGAGTTAACACGGTAACT 67481

Qy      1243 ------------------------------------------------------------ 1243

Db     67480 GGAGAGGGGCTCTTATGACAAGGAGAAAGGAAATTTTTTAAAGGGTGCATGTGAAAGTTA 67421

Qy      1243 ------------------------------------------------------------ 1243

Db     67420 TCAGGTTTGAGGACTTGGGGATGACTTATCAGACCGGGGGATGAGACGTTTAGATCAAAA 67361

Qy      1243 ------------------------------------------------------------ 1243

Db     67360 AAACCCTTTTGAATAGAAGCCAAATTACCTAAAAGTTAATTATCTACTGAAACTGTAACT 67301

Qy      1243 ------------------------------------------------------------ 1243

Db     67300 TTTCCTATACTATTACGGGAACTAAAATCCAATAGAGAAATCCCGAGATTCTCTTTTGGT 67241

Qy      1243 ------------------------------------------------------------ 1243

Db     67240 CTTCATGTGGTTGGTGGGCAGAGGTAAGGCAGAGGCTTAGAATATGAAAAAACATTTAAC 67181

Qy      1243 ------------------------------------------------------------ 1243

Db     67180 AAAATAATAAAGAATGAAACATAAAAATGTTAAGATCATTCTAAAATATTGAGGATCCTG 67121

Qy      1243 ------------------------------------------------------------ 1243

Db     67120 CAGACTTTTTTGGCAGCAAGTTTAAAATGACCCCGGTTATAATTATCTGCAGTGAAACTC 67061

Qy      1243 ------------------------------------------------------------ 1243

Db     67060 AGTATTTCAGGAACCGGCTTAAAGGAGGTAGAGAGTTGGGCTTATCCTGCATTTGTACAC 67001

Qy      1243 ------------------------------------------------------------ 1243

Db     67000 AGCCATGTGACGAGAGGTCACTAGTGCCAAAGCATTGTTCATTTGGGAAAGAGAGTGTTT 66941

Qy      1243 ------------------------------------------------------------ 1243

Db     66940 GTAGTTATGATCTTAGTATACATACTTACAGGGAGTAAAGTGAAGGCAAGAAAGGACTGA 66881

Qy      1243 ------------------------------------------------------------ 1243

Db     66880 TAATTGGAGGGCACAGAAGCAGCAATGGTCTTATATAGATCAAACTTGAGCCAAATGAAC 66821

Qy      1243 ------------------------------------------------------------ 1243

Db     66820 ATTAGTTATATTTGGAGAAATTGTGAAAATGAGAGGGGCAAGATAGGATGTTTGGGTCAG 66761

Qy      1243 ------------------------------------------------------------ 1243

Db     66760 ATAGGACACTGAAATTTGGAAAATGAAATCAATATGTTTTGTTTTCAAACATCTAGAGTA 66701

Qy      1243 ------------------------------------------------------------ 1243

Db     66700 GAAAATGAATATCTGGTTTGGCCTCGCTGGAGATGTTTCACCCCTTGAGTGTGACTAGGA 66641

Qy      1243 ------------------------------------------------------------ 1243

Db     66640 CAAACCACAGTCATGTATAGGGGTTTATTCACTTCACAGGTTCTTGTCTAATCTATGTAA 66581

Qy      1243 ------------------------------------------------------------ 1243

Db     66580 CATTGAAATTCTGCCTTCCTTACCACACCCATATTCATCAAGCACATTTGCTACTTTCAA 66521

Qy      1243 ------------------------------------------------------------ 1243

Db     66520 GTTTTTCTACCTCAGCATTCTCCCACAACCTCCTTCAGGAGTAAAGACAGCCATGTTTTG 66461

Qy      1244 ---------AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGl 1260
                      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66460 TGCCTTTGTAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGA 66401

Qy      1260 uIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAl 1280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66400 GATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGC 66341

Qy      1280 aGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSe 1300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66340 GGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTC 66281

Qy      1300 rPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSe 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66280 GTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTC 66221

Qy      1320 rPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAs 1340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66220 ATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGA 66161

Qy      1340 pAlaGlyValArgCys-------------------------------------------- 1345
             ||||||||||||||||
Db     66160 TGCTGGCGTGAGGTGCTCTGGTAAGTGTGAAAAGGCTGCACTTAAACTGAGCCTCCCACT 66101

Qy      1345 ------------------------------------------------------------ 1345

Db     66100 AATGGTAGAGAACACTGGGAAACAGGCTCACTCAGAGAGTTGAGCTGAAGATACTAGGAT 66041

Qy      1345 ------------------------------------------------------------ 1345

Db     66040 TTCTTGTTGAAGAATGAAAGGCTATGTTAAACTAACGTTTTTCCACAACTAAATAAGATA 65981

Qy      1345 ------------------------------------------------------------ 1345

Db     65980 ATGTTCCACTGTTATCTAAAATAGCCCCCATAATTTTGAGGGAAGAAAGGTGGATTTGGC 65921

Qy      1345 ------------------------------------------------------------ 1345

Db     65920 CTTAGACTTTGAAGCAAGAGGATTCTGACTGATATTTTAAATGACATCCAGTATTCACAG 65861

Qy      1345 ------------------------------------------------------------ 1345

Db     65860 GAGATGTTAAAAGGATATTTAAATGACATCCGATAATCGCAGGAGATGTAAAAAGAAATA 65801

Qy      1346 ------------------------------------------------SerGlyGlnSerL 1350
                                                             |||||||||||||
Db     65800 CCCAGTTCCTTTTCGTTCATTTACGTCTCTCTCTCTCTCTCTCTCTCAGGACAGTCGC 65741

Qy      1350 euLysSerLeuAsnAlaSerSer------------------------------------- 1357
             |||||||||||||||||||||||
```

```
Db   65740 TGAAATCACTGAATGCCTCCTCAGGTATAGTCTTCTGCCTATCAGTGCCGACACAAATAC 65681
Qy    1357 ------------------------------------------------------------ 1357
Db   65680 CTAAGGCTACATGGAAGGTCAAGGGAGGGCAGAAACTGCTATGATAATCAAAGAAGACTG 65621
Qy    1357 ------------------------------------------------------------ 1357
Db   65620 GGGACTGCAGAAATATGAGTCTGTGGGTCGCAGCCAGTGGTATCAATAGCTCTGTGACTG 65561
Qy    1357 ------------------------------------------------------------ 1357
Db   65560 CTTTGCTATGCAGTCCTTCTCTCTAGAAAAATCCTTTTTACATGCAGAGAAGGATAAGAC 65501
Qy    1357 ------------------------------------------------------------ 1357
Db   65500 TCTCATGGTACCCTCATAGATAATCTTTCTAGATTCTCAAATATTCTTTATGAGACTGGC 65441
Qy    1357 ------------------------------------------------------------ 1357
Db   65440 AGAAACAGCTCAAGCTTCAGGAGAATCACTTAAACCCAGGAAGCAGAGATTGCAGTGAGC 65381
Qy    1357 ------------------------------------------------------------ 1357
Db   65380 CGAGATCGTGCCAGTACACTCAGCCTGGGTGACAGAATGAGACTCCATCTCAAAAAAAAA 65321
Qy    1357 ------------------------------------------------------------ 1357
Db   65320 AAAAAAAGAAAGAAACAGCTCAGCTCAAGCCCTAGTATTTTCTGATTCTGGTGTTTTAGT 65261
Qy    1357 ------------------------------------------------------------ 1357
Db   65260 AGTGAATGCAATGCAAGAGGAAAGGTTTAGTTATCAAATCTCTTGTATAAAATTGGAAGT 65201
Qy    1357 ------------------------------------------------------------ 1357
Db   65200 AATTTAGGGCAATAGAATTATGGAAAGATTTGGATTCAGTTGTTCCAAGATAGTTTTCAT 65141
Qy    1357 ------------------------------------------------------------ 1357
Db   65140 CTAAAATAGCTCTCAAAGATTACGTACTGTTCATGAATATGTTGAAATGTATAGCCATTA 65081
Qy    1357 ------------------------------------------------------------ 1357
Db   65080 TTCTTCTCAGCAACTAGGGAACCAGTTTTTTTACAATGTATAAACCCACAATTGATATC 65021
Qy    1358 --------------------------------------GlyHisLeuAlaLeuIleLeuS 1365
                                                 ||||||||||||||||||||||
Db   65020 TACTAGTTTAATACTTATCAATGCTTAATTTCTATACAGGTCATTTAGCACTTATTTTAT 64961
Qy    1365 erSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArgV 1385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64960 CCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAG 64901
Qy    1385 alGlnLysGlnLysHisLeuProLeuArgVal---------------------------- 1395
           ||||||||||||||||||||||||||||||||
Db   64900 TTCAGAAACAAAAACATCTGCCCCTCAGAGG-TGAGAGCCCCACTAGGATTGTATAAAAT 64842
Qy    1395 ------------------------------------------------------------ 1395
Db   64841 GTGAGTTACCTGCCACCAAGGGCACACAGGATGGGGGATTCTTGCAGAATGGAAGAAATG 64782
Qy    1395 ------------------------------------------------------------ 1395
Db   64781 CACTTGCCACTTTGAGTGAGATGGATTGGAAAGTGCCAGGCCAAATGCAGAAACATGTGG 64722
Qy    1395 ------------------------------------------------------------ 1395
Db   64721 AATGATCAAGAATCAGCCAATCCACAATTTCCATTTACACAAATGTAATCTTTTAGAAGC 64662
Qy    1395 ------------------------------------------------------------ 1395
Db   64661 TGTAGCTTGCCATACAGACGATGAGTCAGATAGGCAAATATCTTAGCTGATAGAAGGAAG 64602

Qy    1395 ------------------------------------------------------------ 1395
Db   64601 GTAATTGTATGTTTTGGTTCAGAATCTAGACTGCAGAAATAGCCAATAGAAAATGTAGAA 64542
Qy    1395 ------------------------------------------------------------ 1395
Db   64541 GCATGGTGGAGGATTAAGAACAAGAGTTTCCTGGGTTTGAACACCACTTATATAACTTAC 64482
Qy    1395 ------------------------------------------------------------ 1395
Db   64481 TAGCTGTGTGGCCTTTGGCTAGAAACTACCTTTCCAAACTCCAATTTCCTAATTTGTAAA 64422
Qy    1395 ------------------------------------------------------------ 1395
Db   64421 GATAATAATAGTCTTACCTCATCAGGTTAAGATAGTTAAATAAATTAATCATCTGTATGA 64362
Qy    1395 ------------------------------------------------------------ 1395
Db   64361 AAAATGGCCAGCACATAGTAAATACTATGTAAAGAACTTGGTAATAAAAGTGGAAGAATA 64302
Qy    1395 ------------------------------------------------------------ 1395
Db   64301 TCAAGGGAGAAAAAGGAGGATAAATTTTAAAAATGTAAAAAATTCACTAGCAGGGGTAGG 64242
Qy    1395 ------------------------------------------------------------ 1395
Db   64241 TGCTGAAAGAGCCATTGGATTTTAGCTCAGAGGGATTAATATGTGTCAAACCTTAGCATT 64182
Qy    1395 ------------------------------------------------------------ 1395
Db   64181 AAGTGTCCTAATATAAGGGATAGGCCATTTGAATTTGAATCCTAACCATGAAGCATCCAT 64122
Qy    1396 --------SerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluT 1413
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64121 TCTCAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGA 64062
Qy    1413 hrCysLeuLysArgGluAspProHisGlyThrArgThrSer 1426
           |||||||||||||||||||||||||||||||||||||||||
Db   64061 CCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCA 64021
```

Search completed: May 12, 2003, 19:20:08
Job time : 6250 secs

THIS PAGE BLANK (USPTO)

```
OM protein - protein search, using sw model

Run on:          May 12, 2003, 06:24:52 ; Search time 52 Seconds
                                          (without alignments)
                                          5757.432 Million cell updates/sec

Title:           US-09-759-130B-381
Perfect score:   8138
Sequence:        1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:      671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_21:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8109	99.6	1453	4	Q9NR16	Q9nr16 homo sapien
2	3441.5	42.3	1151	4	Q07899	Q07899 homo sapien
3	3429	42.1	1116	4	Q07898	Q07898 homo sapien
4	3429	42.1	1156	4	Q07900	Q07900 homo sapien
5	3396.5	41.7	1149	4	Q07901	Q07901 homo sapien
6	3272	40.2	1121	11	Q99MX8	Q99mx8 mus musculu
7	2990	36.7	2403	4	Q9UGM2	Q9ugm2 homo sapien
8	2987.5	36.7	2426	4	Q9UGM3	Q9ugm3 homo sapien
9	2987	36.7	2412	4	Q9UJ57	Q9uj57 homo sapien
10	2982	36.6	2413	4	Q96DU4	Q96du4 homo sapien
11	2980	36.6	2413	4	Q9UKJ4	Q9ukj4 homo sapien
12	2448	30.1	1785	4	Q9Y211	Q9y211 homo sapien
13	2439.5	30.0	2153	5	O97375	O97375 strongyloce
14	2435	29.9	1785	4	Q9Y4V9	Q9y4v9 homo sapien
15	2246.5	27.6	2043	5	O96943	O96943 geodia cydo
16	2156.5	26.5	1594	6	Q95218	Q95218 oryctolagus
17	1962.5	24.1	2083	11	Q60997	Q60997 mus musculu
18	1845	22.7	1957	11	Q9JMJ9	Q9jmj9 mus musculu
19	1594.5	19.6	1188	5	Q9NC87	Q9nc87 strongyloce
20	1586	19.5	710	6	Q95JE5	Q95je5 ovis aries
21	1431.5	17.6	749	5	Q9NC88	Q9nc88 strongyloce
22	1399	17.2	804	6	Q29113	Q29113 sus scrofa
23	1341	16.5	546	6	Q29111	Q29111 sus scrofa
24	1301	16.0	1075	5	Q9NC90	Q9nc90 strongyloce
25	1277	15.7	973	5	Q9NC89	Q9nc89 strongyloce
26	1219	15.0	369	6	Q29110	Q29110 sus scrofa
27	1205	14.8	420	6	Q95JD9	Q95jd9 ovis aries
28	1085.5	13.3	480	6	Q29112	Q29112 sus scrofa
29	1078.5	13.3	600	6	Q28910	Q28910 bos taurus
30	1048.5	12.9	1290	11	Q62827	Q62827 rattus norv
31	1041.5	12.8	470	13	Q98TD1	Q98td1 gallus gall
32	968.5	11.9	575	4	Q8WTU2	Q8wtu2 homo sapien
33	866.5	10.6	1036	5	O97378	O97378 strongyloce
34	763.5	9.4	437	6	Q29109	Q29109 sus scrofa
35	723.5	8.9	749	11	Q99JC8	Q99jc8 rattus norv
36	703	8.6	209	6	Q28881	Q28881 ovis aries
37	686.5	8.4	457	5	Q9NC91	Q9nc91 strongyloce
38	665	8.2	918	13	Q92098	Q92098 petromyzon
39	664.5	8.2	268	6	Q9TV20	Q9tv20 oryctolagus
40	664.5	8.2	277	6	Q9TV21	Q9tv21 oryctolagus
41	662.5	8.1	264	6	Q9TV22	Q9tv22 oryctolagus
42	650	8.0	505	4	Q96I80	Q96i80 homo sapien
43	632	7.8	754	11	Q91VN8	Q91vn8 mus musculu
44	620.5	7.6	665	11	Q91WN5	Q91wn5 mus musculu
45	603.5	7.4	592	4	Q8WWJ7	Q8wwj7 homo sapien

ALIGNMENTS

```
RESULT 1
Q9NR16
ID   Q9NR16      PRELIMINARY;      PRT;  1453 AA.
AC   Q9NR16;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Scavenger receptor cysteine-rich type 1 protein M160 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20540099; PubMed=11086079;
RA   Gronlund J., Vitved L., Lausen M., Skjodt K., Holmskov U.;
RT   "Cloning of a Novel Scavenger Receptor Cysteine-Rich Type I
RT   Transmembrane Molecule (M160) Expressed by Human Macrophages.";
RL   J. Immunol. 165:6406-6415(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Groenlund J., Vitved L., Lausen M., Skjodt K., Holmskov U.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF264014; AAF91396.1; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 12.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 11.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_5.
DR   PROSITE; PS50287; SRCR_2; 12.
KW   Receptor; Signal.
FT   SIGNAL        1     42       POTENTIAL.
SQ   SEQUENCE   1453 AA;  159263 MW;  15445CE9868CF58A CRC64;

  Query Match             99.6%;  Score 8109;  DB 4;  Length 1453;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1449;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60

Qy       61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120

Qy      121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180

Qy      181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240

Qy      241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300

Qy      301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360

Qy      361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420

Qy      421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480

Qy      481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540

Qy      541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600

Qy      601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660

Qy      661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720

Qy      721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780

Qy      781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840

Qy      841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900

Qy      901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960

Qy      961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020

Qy     1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080

Qy     1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140
            ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     1081 VAFNATVSAHFGEGSGPIWLDDLNCTGMESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140

Qy     1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200

Qy     1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260
            ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSTAEETWITCEDRIRVRGGDTECSGRVE 1260

Qy     1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320

Qy     1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380
            |||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||||
Db     1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGRLALILSSIFGLLLPVLFILFL 1380

Qy     1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440

Qy     1441 SLLGVLPASEATK 1453
            |||||||||||||
Db     1441 SLLGVLPASEATK 1453
```

```
RESULT 2
Q07899
ID   Q07899     PRELIMINARY;     PRT;  1151 AA.
AC   Q07899;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   M130 antigen cytoplasmic variant 1 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93380506; PubMed=8370408;
RA   Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA   Mason D.Y.;
RT   "A new macrophage differentiation antigen which is a member of the
RT   scavenger receptor superfamily.";
RL   Eur. J. Immunol. 23:2320-2325(1993).
DR   EMBL; Z22969; CAA80542.1; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 9.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 9.
DR   PROSITE; PS00420; SRCR_1; 4.
DR   PROSITE; PS50287; SRCR_2; 9.
KW   Signal.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41   1151       M130 ANTIGEN CYTOPLASMIC VARIANT 1.
SQ   SEQUENCE   1151 AA;  124820 MW;  3CD0BAD857B6EB91 CRC64;

  Query Match             42.3%;  Score 3441.5;  DB 4;  Length 1151;
  Best Local Similarity   55.2%;  Pred. No. 4e-260;
  Matches  608;  Conservative  166;  Mismatches  311;  Indels   17;  Gaps    6;

Qy      364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
            | | |||| || | |||||||:: |:| |:|   |  |   |:| ||||| ::     |
Db       41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy      424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
             | :   ||:: :|| ||||||||| :||  |:  |  :  |||| ||| ::|::||
Db      101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy      483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
            : | ||:|:|:|| ||||| | ::  :|:|:|:|| ||  :   | : | | ||||| ||
```

```
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
          : | |||| :|:|:| ||||||| | ||  | ||  |    |||| |    ||||||| |||
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
          |||:|||||:|  ||| | || ||:::  :|  ||| |:| |||| ||| | |   :| |
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
          ::  || : |:|:|| || ||| ||:|||| || ||||| |||  :|  :|  :|   |||:
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
          |:|||||| ||||:: | :  :    :   :   :  | | | | ||||| |:|      |
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
          || : |||||:||||| |:|||||||||| ||| |:||||||| ||:|||||| ||    :
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
          |:  | |||:|||  |||:||||| |:||:|||:     || || |||:|||||||||::||
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
          ||||: |:|:||:  || |||||::||| ||| |||:|| |||| |||  |  :  |: : ::
Db    581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
          | | ||| | |    : :| :| |||  |       ||||:|: :| |  | ::   |     :
Db    641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
          || ||: |:|   :||||:| ||||||||||:| ||||||| ||||||||||::|||| |
Db    701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
          ||| |||||||:||||||:: | | || :|||  | ||||  :||||||||||||||| :||
Db    761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
          | ||    |:|||||||||||| ||:|:  :::      |:|||||||| : | :: | | | |
Db    821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
           ||||::||||    ::|||  |:|||:|::||||||||:::||:: |  | ||||||||
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
          | ||||||||||||| :|:||||||||| || | ::| ||||| |||:::::|||||| |
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
          ||| |: || |:||||||| | |:  |::         :   ||   : |   | |::|| :|
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGC 1434
          :   |||  | |:|:    :|:  :|   :  :  :  :| :||  :   :  :|:  :  :|
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADDLDLMNSSEN--SHES 1115

Qy   1435 EDASDTSLLGV---LPASEATK 1453
          | |    |: |     || |    |
Db   1116 ADFSAAELISVSKFLPISGMEK 1137
```

RESULT 3
Q07898
ID   Q07898   PRELIMINARY;   PRT; 1116 AA.
AC   Q07898;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   M130 antigen precursor.
GN   CD163.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93380506; PubMed=8370408;
RA   Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA   Mason D.Y.;
RT   "A new macrophage differentiation antigen which is a member of the
RT   scavenger receptor superfamily.";
RL   Eur. J. Immunol. 23:2320-2325(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ritter M., Buechler C., Langmann T., Schmitz G.;
RT   "Genomic organization of the human CD163 gene.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z22968; CAA80541.1; -.
DR   EMBL; Y18388; CAB45233.1; -.
DR   EMBL; Y18389; CAB45233.1; JOINED.
DR   EMBL; Y18390; CAB45233.1; JOINED.
DR   EMBL; Y18391; CAB45233.1; JOINED.
DR   EMBL; Y18392; CAB45233.1; JOINED.
DR   EMBL; Y18393; CAB45233.1; JOINED.
DR   EMBL; Y18394; CAB45233.1; JOINED.
DR   EMBL; Y18395; CAB45233.1; JOINED.
DR   EMBL; Y18396; CAB45233.1; JOINED.
DR   EMBL; Y18397; CAB45233.1; JOINED.
DR   EMBL; Y18398; CAB45233.1; JOINED.
DR   EMBL; Y18399; CAB45233.1; JOINED.
DR   EMBL; Y18400; CAB45233.1; JOINED.
DR   EMBL; Y18401; CAB45233.1; JOINED.
DR   EMBL; Y18402; CAB45233.1; JOINED.
DR   EMBL; Y18403; CAB45233.1; JOINED.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 9.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 9.
DR   PROSITE; PS00420; SRCR_1; 4.
DR   PROSITE; PS50287; SRCR_2; 9.
KW   Antigen; Signal.
FT   SIGNAL   1   40   POTENTIAL.
FT   CHAIN   41   1116   M130 ANTIGEN.
SQ   SEQUENCE   1116 AA; 120979 MW; E805BC80B52E4009 CRC64;

Query Match   42.1%; Score 3429; DB 4; Length 1116;
Best Local Similarity   56.2%; Pred. No. 3.7e-259;
Matches 598; Conservative 161; Mismatches 294; Indels   12; Gaps   4;

```
Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
          | | |||| || | |||||||:: |:| |:|:  |  |   |:| |||| ::        |
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
           |  :   ||:: :|| ||||||||| :||  | : |   : |||| ||| ::|::||
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
          : | ||:|:|:|| ||||| | ::  :|:|:|:|| ||  :    | : | | | |||||| ||
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
          : | |||| :|:|:| ||||||| | ||  | ||  |    ||||     ||||||| |||
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
          |||:|||||:|  ||| | || ||:::  :|  ||| |:| |||| ||| | |   :| |
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
```

```
              :: || : |:|:|| || ||| ||:|||| || ||||| ||| :|  :| :|  |||:
Db      341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy      723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
             |:||||| ||||:: | : :    :   :   :  | | | | ||||| |:|     |
Db      401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy      783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
              || : |||||:||||| |:|||||||||| ||| |:||||||| ||:|||||| ||   :
Db      461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy      843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
            |:  | |||:||| |||:||||| |:||:|||:   || || |||:|||||||||::||
Db      521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy      903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
            ||||: |:|:||:  || |||||::||| ||| |||:|| || |||| ||     |: : ::
Db      581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy      963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
            | | ||| | |   : :| :| ||| |          ||||:|: :| | | ::    |      :
Db      641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy     1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
            || ||: |:|  :||||:| ||||||||||:| |||||| |||||||||||:::|||| |
Db      701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy     1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
            ||| |||||||:||||||:: | | || :||| | |||| :|||||||||||||| :||
Db      761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy     1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
            | ||    |:|||||||||||| ||:||: :::      |:||||||| : | :: | |  |
Db      821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy     1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
             ||||::||||     ::||| |:|||:|::||:|||||||:::||:: | | ||||||||
Db      881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy     1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
            | ||||||||||||| :|:||||||||| | |  | ||||||||| |||:: ||::::|||||| |
Db      941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy     1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
            ||| |: || |:||||||| | |: . |::     :   ||   : |    | |::|| :|
Db     1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy     1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
            :    ||| | |:|:   | ||:|   : :  : || :|| :|
Db     1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102
```

RESULT 4
Q07900

```
ID   Q07900      PRELIMINARY;      PRT;  1156 AA.
AC   Q07900;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   M130 antigen cytoplasmic variant 2 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93380506; PubMed=8370408;
RA   Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA   Mason D.Y.;
RT   "A new macrophage differentiation antigen which is a member of the
RT   scavenger receptor superfamily.";
RL   Eur. J. Immunol. 23:2320-2325(1993).
DR   EMBL; Z22970; CAA80543.1; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 9.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 9.
DR   PROSITE; PS00420; SRCR_1; 4.
DR   PROSITE; PS50287; SRCR_2; 9.
KW   Signal.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41   1156       M130 ANTIGEN CYTOPLASMIC VARIANT 2.
SQ   SEQUENCE   1156 AA;  125352 MW;  3BEEC22D805B1D5C CRC64;

  Query Match             42.1%;  Score 3429;  DB 4;  Length 1156;
  Best Local Similarity   56.2%;  Pred. No. 3.8e-259;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels   12;  Gaps    4;

Qy      364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
            | | |||| || |  |||||||:: |:| |:|:  |   |    |:| ||||| ::     |
Db       41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy      424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
              |  :  ||:: :|| |||||||| :||  | : |  : |||| ||| ::|::||
Db      101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy      483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
             : |:||:|:|:||  |  :|:::|:|:|| ||   :   : |  | | |||| ||
Db      161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy      543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
              : | |||| :|:|:| ||||||| | ||  | | ||  |   ||||  ||||||| |||
Db      221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy      603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
             |||:||||:| ||| | || |||:::  :|  ||| |:| |||| ||| | |   :| |
Db      281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy      663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
            :: || : |:|:|| || ||:| ||:|||| || ||||| ||| :|  :| :|  |||:
Db      341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy      723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
             |:||||| ||||:: | : :    :   :   :  | | | | ||||| |:|     |
Db      401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy      783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
              || : |||||:||||| |:|||||||||| ||| |:||||||| ||:|||||| ||   :
Db      461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy      843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
            |:  | |||:||| |||:||||| |:||:|||:   || || |||:|||||||||::||
Db      521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy      903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
            ||||: |:|:||:  || |||||::||| ||| |||:|| || |||| ||     |: : ::
Db      581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy      963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
            | | ||| | |   : :| :| ||| |          ||||:|: :| | | ::    |      :
Db      641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy     1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
            || ||: |:|  :||||:| ||||||||||:| |||||| |||||||||||:::|||| |
Db      701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy     1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
            ||| |||||||:||||||:: | | || :||| | |||| :|||||||||||||| :||
Db      761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy     1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
            | ||    |:|||||||||||| ||:||: :::      |:||||||| : | :: | |  |
Db      821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy     1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
             ||||::||||     ::||| |:|||:|::||:|||||||:::||:: | | ||||||||
```

```
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
          | |||||||||||| :|:||||||||| || | ::| |||||| |||:::::|||||| |
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
          ||| |: || |:||||||| | |:  |::        :    ||    : |    | |::|| :|
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
          :   |||  | |:|:   | ||:|    :  :  :  || :|| :|
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102
```

RESULT 5
Q07901

```
ID   Q07901      PRELIMINARY;      PRT;  1149 AA.
AC   Q07901;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   M130 antigen extracellular variant precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93380506; PubMed=8370408;
RA   Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA   Mason D.Y.;
RT   "A new macrophage differentiation antigen which is a member of the
RT   scavenger receptor superfamily.";
RL   Eur. J. Immunol. 23:2320-2325(1993).
DR   EMBL; Z22971; CAA80544.1; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 9.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 9.
DR   PROSITE; PS00420; SRCR_1; 4.
DR   PROSITE; PS50287; SRCR_2; 9.
KW   Signal.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41   1149       M130 ANTIGEN EXTRACELLULAR VARIANT.
SQ   SEQUENCE   1149 AA;  124328 MW;  FDBA6D85CF97BB60 CRC64;

  Query Match             41.7%;  Score 3396.5;  DB 4;  Length 1149;
  Best Local Similarity   54.4%;  Pred. No. 1.3e-256;
  Matches  597;  Conservative  161;  Mismatches  295;  Indels   45;  Gaps    5;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
          | | |||| || | |||||||:: |:| |:|:  |  |    |:| ||||| ::    |
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
           |  :  ||::  :|| |||||||| :||  |  : |  :  ||||| ||| ::|::||
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
          : | ||:|:|:|| ||||| | ::  :|:|:|:|| ||  :    | : | | ||||| ||
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
          : | |||| :|:|:| ||||||| | || | ||  |    |||| |   ||||||| |||
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
           |||:|||||:| ||| | || ||::: :| ||| |:| |||| ||| | |  :| |
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
          :: || : |:|:|| || ||| ||:|||| || ||||| ||| :| :| :| |||:
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
          |:|||||| ||||:: | : :   :  :  :  | | | | ||||| |:|   |
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
           || : |||||:||||| |:|||||||||| ||| |:||||||| ||:|||||| ||  :
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSR------ 896
          |: | |||:||| |||:|||||.|:||:|||:  || || |||:||||||
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSSKTQKTS 580

Qy    897 --------------------------YTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHW 929
                                    ||::|||||: |:|:||:  || ||||::||
Db    581 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHW 640

Qy    930 DPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPC 989
          | ||| |||:|| || |||| ||   |: : ::| | ||| |  | : :| :| ||| |
Db    641 DIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLC 700

Qy    990 IHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRV 1049
                 ||||:|:|:| |  | ::   |       :|| ||: |:|  :||||:| ||||||
Db    701 PSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRV 760

Qy   1050 EIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTE 1109
          ||||:| ||||||| ||||||||||::||||| |||||||||||:||||||:: | |
Db    761 EIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKE 820

Qy   1110 SHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGR 1169
          | :||| | |||| :||||||||||||||:|:||| ||   |:|||||||||| ||:|:|:
Db    821 SRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGK 880

Qy   1170 RNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWER 1229
           :::    |:|||||| : | :: | | |  | ||||::|||   ::||| |:|||:
Db    881 SSMSETTVGVVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEK 940

Qy   1230 RISSPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGC 1289
          |::||:|||||||:::||:: | | ||||||||| |||||||||||| :|:||||||||
Db    941 RLASPSEETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGC 1000

Qy   1290 GSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQS 1349
          | || | ::| ||||| |||::::|||||||| ||| |: || |:||||||| | |: |
Db   1001 GPALKAFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDIS 1060

Qy   1350 LK------SLNASSGHLALILSSIFGLLLLVLFI--LFLTWCRVQKQHLPLRVSTRRRG 1401
          ::      :    ||    : |    | |::|| :|:   |||  | |:|:   | ||:|
Db   1061 VQKTPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQR---LAVSSRGEN 1117

Qy   1402 SLEENLFHEMETCLKRED 1419
           :  :  :  || :|| :|
Db   1118 LVHQIQYREMNSCLNADD 1135
```

RESULT 6
Q99MX8

```
ID   Q99MX8      PRELIMINARY;      PRT;  1121 AA.
AC   Q99MX8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Macrophage hemoglobin scavenger receptor CD163 precursor.
GN   CD163.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21243729; PubMed=11345593;
RA   Schaer D.J., Boretti F.S., Hongegger A., Poehler D., Linnscheid P.,
RA   Staege H., Muller C., Schoedon G., Schaffner A.;
```

```
RT   "Molecular cloning and characterization of the mouse CD163 homologue,
RT   a highly glucocorticoid-inducible member of the scavenger receptor
RT   cysteine-rich family.";
RL   Immunogenetics 53:170-177(2001).
DR   EMBL; AF274883; AAK16065.1; -.
DR   MGD; MGI:2135946; Cd163.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 9.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 9.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_3.
DR   PROSITE; PS50287; SRCR_2; 9.
KW   Receptor; Signal.
FT   SIGNAL        1     38       POTENTIAL.
SQ   SEQUENCE   1121 AA;  120889 MW;  04428DF18AD26187 CRC64;

  Query Match            40.2%;  Score 3272;  DB 11;  Length 1121;
  Best Local Similarity  52.4%;  Pred. No. 6.9e-247;
  Matches  582;  Conservative  174;  Mismatches  332;  Indels   22;  Gaps    9;

Qy    320 GSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADL--ELRLADGSNN 377
Db      3 GHRMVLLGGAGSPGCKRFV----HLGFFVVAVSSLLSASAVTNAPGEMKKELRLAGGENN 58

Qy    378 CSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSIS 437
Db     59 CSGRVELKIHDKWGTVCSNGWSMNEVSVVCQQLGCPTSIKALGWANSSAGSGYIWMDKVS 118

Qy    438 CTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLV-GAHSPCYGRLEVKYQGE 496
Db    119 CTGNESALWDCKHDGWGKHNCTHEKDAGVTCSDGSNLEMRLVNSAGHRCLGRVEIKFQGK 178

Qy    497 WGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCE 556
Db    179 WGTVCDDNFSKDHASVICKQLGCGSAISFSGSAKLGAGSGPIWLDDLACNGNESALWDCK 238

Qy    557 HSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKA 616
Db    239 HRGWGKHNCDHAEDVGVICLEGADLSLRLVDGVSRCSGRLEVRFQGEWGTVCDDNWDLRD 298

Qy    617 AAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSE 676
Db    299 ASVVCKQLGCPTAISAIGRVNASEGSGQIWLDNISCEGHEATLWECKHQEWGKHYCHHRE 358

Qy    677 DVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSA 736
Db    359 DAGVTCSDGADLELRLVGGGSRCAGIVEVEIQKLTGKMCSRGWTLADADVVCRQLGCGSA 418

Qy    737 IRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPR 796
Db    419 LQTQAKIYSKTGATNTWLFPGSCNGNETTFWQCKNWQWGGLSCDNFEEAKVTCSGHREPR 478

Qy    797 LVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGL 856
Db    479 LVGGEIPCSGRVEVKHGDVWGSVCDFDLSLEAASVVCRELQCGTVVSILGGAHFGEGSGQ 538

Qy    857 TWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEIN 916
Db    539 IWGEEFQCSGDESHLSLCSVAPPLDRTCTHSRDVSVVCSRYIDIRLAGGESSCEGRVELK 598

Qy    917 VLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLL 976
Db    599 TLGAWGPLCSSHWDMEDAHVLCQQLKCGVAQSIPEGAHFGKGAGQVWSHMFHCTGTEEHI 658

Qy    977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPC--LANVSDPYLSAVPEGSALICLEDK 1034
Db    659 GDCLMTALGAPTCSEGQVASVICSGNQSQTLLPCSSLSPVQTT-SSTIPKESEVPCIASG 717

Qy   1035 RLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEG 1094
Db    718 QLRLVGGGGRCAGRVEVCHEGSWGTVCDDNWDMTDANVVCKQLDCGVAINATGSAYFGEG 777

Qy   1095 SGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAG 1154
Db    778 AGAIWLDEVICTGKESHIWQCHSHGWGRHNCRHKEDAGVICSEFMSLRLTNEAHKENCTG 837

Qy   1155 RLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPK 1214
Db    838 RLEVFYNGTWGSIGSSNMSPTTVGVVCRQLGCADNGTVKPIPSDKTPSRPMWVDRVQCPK 897

Qy   1215 THISIWQCLSAPWERRISSP-AEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDD 1273
Db    898 GVDTLWQCPSSPWKQRQASPSSQESWIICDNKIRLQEGHTDCSGRVEIWHKGSWGTVCDD 957

Qy   1274 SWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQS 1333
Db    958 SWDLNDAKVVCKQLGCGQAVKALKEAAFGPGTGPIWLNEIKCRGNESSLWDCPAKPWSHS 1017

Qy   1334 DCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFLTWC----RVQKQK 1389
Db   1018 DCGHKEDASIQCLPKMTSESHHGTGHPTLTALLVCGAILLVLLIVFLLWTLKRRRIQR-- 1075

Qy   1390 HLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
Db   1076 ---LTVSSRGEVLIHQVQYQEMDS--KADD 1100

RESULT 7
Q9UGM2
ID   Q9UGM2      PRELIMINARY;      PRT;  2403 AA.
AC   Q9UGM2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DMBT1/8kb.1 protein.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE=20065089; PubMed=10597221;
RA   Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herbertz S.,
RA   Madsen J., Kioschis P., Coy J.F., Poustka A.;
RT   "The genomic structure of the DMBT1 gene: evidence for a region with
RT   susceptibility to genomic instability.";
RL   Oncogene 18:6233-6240(1999).
CC   -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR   EMBL; AJ243224; CAB63942.1; -.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 14.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00042; CUB; 2.
DR   SMART; SM00202; SR; 14.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS01180; CUB; 2.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR   PROSITE; PS50287; SRCR_2; 14.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
FT   CHAIN        26   2403       DMBT1/8KB.1 PROTEIN.
SQ   SEQUENCE   2403 AA;  259573 MW;  902622DE7399AD38 CRC64;

  Query Match            36.7%;  Score 2990;  DB 4;  Length 2403;
  Best Local Similarity  40.5%;  Pred. No. 2e-224;
  Matches  619;  Conservative  177;  Mismatches  504;  Indels  230;  Gaps   23;

Qy     44 TDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM---FRF 100
Db    230 SSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQF 289
```

```
Qy    101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG---------- 149
Db    290 GQG---SGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICSAPLSRPTPSPDT 346

Qy    150 ----------EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGC 198
Db    347 WPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGC 406

Qy    199 PSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDS- 257
Db    407 GWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSDTL 466

Qy    258 ------------SDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLG 304
Db    467 PTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLG 526

Qy    305 CGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS-D 363
Db    527 CGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHNGWLSHNCGHSEDAGVICSGP 586

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db    587 ESSLALGLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNAR 646

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICS-------------- 469
Db    647 FGQGSGPIVLDDVRCSGHESYLWSCPNNGWLSHNCGHHEDAGVICSAAQSRSTPRPDTLS 706

Qy    470 ---------DKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGC 519
Db    707 TITLPPSTVGSESSLTLRLVNGSDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGC 766

Qy    520 GKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCS--- 576
Db    767 GWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSVSQ 826

Qy    577 -----GDATW-------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAA 617
Db    827 SRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDA 886

Qy    618 AVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
Db    887 NVVCRRLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSED 946

Qy    678 VGVICSDA------------------------SDMELRLVGGSSRCAGKVEVNVQGAVGI 713
Db    947 AGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYQGSWGT 1006

Qy    714 LCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWE 773
Db   1007 VCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNG 1066

Qy    774 WKQTACHLNMEASLICSAHR-QP----------------------RLVGADMPCSGRVE 809
Db   1067 WLSHNCGHSEDAGVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVE 1126

Qy    810 VKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSET 869
Db   1127 VLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHES 1186

Qy    870 HLALCPIVQHPEDTCIHSREVGVVCSRYTD-----------------------VRLVNG 905
Db   1187 YLWSCPHDGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNG 1246

Qy    906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGH 965
Db   1247 GDRCQGRVEVLYRGPWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLD 1306

Qy    966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEG 1025
Db   1307 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPT-PSPDTWPTSHASTAGSE 1365

Qy   1026 SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNA 1085
Db   1366 SSL------ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSA 1419

Qy   1086 TVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC---------- 1135
Db   1420 FGSARFGQGSGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPS 1479

Qy   1136 ------------SEFT-ALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVC 1181
Db   1480 PDTWPTSRASTAGSESTLALRLVN--GGDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVC 1537

Qy   1182 EQLGCGENGVVSLAPLSK---TGSGFMWVDDIQCPKTHISIWQCLSAPW---------- 1227
Db   1538 RQLGC---GWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDA 1594

Qy   1228 ------ERRISSPAEETWIT---------CEDRIRVR---GGDTECSGRVEIWHAGSWGT 1269
Db   1595 GVICSAAQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGD-RCRGRVEVLYRGSWGT 1653

Qy   1270 VCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKP 1329
Db   1654 VCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKG 1713

Qy   1330 WGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
Db   1714 WLTHNCGHHEDAGVICSATQINSTTTDWWH 1743
```

```
RESULT 8
Q9UGM3
ID   Q9UGM3     PRELIMINARY;     PRT;  2426 AA.
AC   Q9UGM3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DMBT1 prototype precursor.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20065089; PubMed=10597221;
RA   Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herbertz S.,
RA   Madsen J., Kioschis P., Coy J.F., Poustka A.;
RT   "The genomic structure of the DMBT1 gene: evidence for a region with
RT   susceptibility to genomic instability.";
RL   Oncogene 18:6233-6240(1999).
CC   -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR   EMBL; AJ243211; CAB63941.1; -.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 14.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00042; CUB; 2.
DR   SMART; SM00202; SR; 14.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS01180; CUB; 2.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR   PROSITE; PS50287; SRCR_2; 14.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW   Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26   2426       DMBT1 PROTOTYPE.
SQ   SEQUENCE   2426 AA;  262052 MW;  5A58FBC076FB7247 CRC64;

  Query Match             36.7%;  Score 2987.5;  DB 4;  Length 2426;
  Best Local Similarity   39.7%;  Pred. No. 3.3e-224;
  Matches  615;  Conservative  181;  Mismatches  498;  Indels  256;  Gaps   22;
```

```
Qy     43 GTD--LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM--- 97
Db     95 GSDSGLALRLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGN 154

Qy     98 FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCY--------- 148
Db    155 AWFGQG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLR 211

Qy    149 ---------------GEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVV 192
Db    212 PESWPVRISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVV 271

Qy    193 CRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTL 252
Db    272 CRQLGCGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGV 331

Qy    253 TCY---------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHK 290
Db    332 ICSAPQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDS 391

Qy    291 WNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFD 350
Db    392 WDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHN 451

Qy    351 CLHQNDVSVICS----------------------DGADLELRLADGSNNCSGRVEVRI 386
Db    452 CQHSEDAGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLY 511

Qy    387 HEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALW 446
Db    512 RGSWGTVCDDSWDTNDANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLW 571

Qy    447 DCTYDGKAKRTCFRRSDAGVICSD-------------KADLDLRLVGAHSPCYGRLEVK 492
Db    572 SCPHNGWLSHNCGHSEDAGVICSDTLPTTTLPASTVGPESSLALRLVNGGDRCQGRVEVL 631

Qy    493 YQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNI 552
Db    632 YRGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYL 691

Qy    553 WDCEHSGWGKHNCVHREDVIVTCS----------------------GDATWGLRLVGG 588
Db    692 WSCPNNGWLSHNCGHHEDAGVICSAAQSRSTPRPDTLSTITLPPSTVGSESSLTLRLVNG 751

Qy    589 SNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLD 648
Db    752 SDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLD 811

Qy    649 DVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICS---------------------DAS 686
Db    812 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSVSQSRPTPSPDTWPTSHASTAGSES 871

Qy    687 DMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFT 746
Db    872 SLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGCGWATSAPGNARFG 931

Qy    747 ERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQ----------- 794
Db    932 QGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAHSWSTPSPDTLPTI 991

Qy    795 --------------PRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840
Db    992 TLPASTVGSESSLALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVVCRQLGCGW 1051

Qy    841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVC----SR 896
Db   1052 AMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAGVICSASQSR 1111

Qy    897 YTD-------------------VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARV 936
Db   1112 PTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANV 1171

Qy    937 LCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVS 996
Db   1172 VCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAG 1231

Qy    997 VICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGF 1056
Db   1232 VICSASQSQPT-PSPDTWPTSHASTAGSESSL------ALRLVNGGDRCQGRVEVLYRGS 1284

Qy   1057 WGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCP 1116
Db   1285 WGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCP 1344

Qy   1117 SRGWGQHDCRHKEDAGVICS----------------------EFTALRLYSETETESC 1152
Db   1345 HNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVN--GGDRC 1402

Qy   1153 AGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSK---TGSGFMWVDD 1209
Db   1403 QGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGC---GWATSAPGNARFGQGSGPIVLDD 1459

Qy   1210 IQCPKTHISIWQCLSAPW----------------ERRISSPAEETWIT-------CEDR 1245
Db   1460 VRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSRASTAGSEST 1519

Qy   1246 IRVR---GGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFG 1302
Db   1520 LALRLVNGGD-RCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQFG 1578

Qy   1303 QGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352
Db   1579 QGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQS 1628
```

```
RESULT 9
Q9UJ57
ID   Q9UJ57      PRELIMINARY;      PRT;  2412 AA.
AC   Q9UJ57;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DMBT1/8kb.2 protein precursor.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TRACHEA;
RX   MEDLINE=99415938; PubMed=10485905;
RA   Holmskov U., Mollenhauer J., Madsen J., Vitved L., Groenlund J.,
RA   Tornoee I., Kliem A., Reid K.B., Poustka A., Skjoedt K.;
RT   "Cloning of gp-340, a putative opsonin receptor for lung surfactant
RT   protein D.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC   -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR   EMBL; AJ243212; CAB56155.1; -.
DR   HSSP; P29392; 1SFP.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 14.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00042; CUB; 2.
DR   SMART; SM00202; SR; 14.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS01180; CUB; 2.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR   PROSITE; PS50287; SRCR_2; 14.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW   Signal.
```

```
FT   SIGNAL          1     25       POTENTIAL.
FT   CHAIN          26   2412       DMBT1/8KB.2 PROTEIN.
SQ   SEQUENCE   2412 AA;  260568 MW;  3F630CCBFFB18EDD CRC64;

  Query Match             36.7%;  Score 2987;  DB 4;  Length 2412;
  Best Local Similarity   40.1%;  Pred. No. 3.5e-224;
  Matches  617;  Conservative  177;  Mismatches  506;  Indels  240;  Gaps   22;

Qy     44 TDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM---FRF 100
          : | |||||   | | ||| ::| |||||| |:|   : |||:|||| :: :     :|
Db    230 SSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQF 289

Qy    101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG---------- 149
          ||     | | |||| | |:|| || | |   | :||| | || || |
Db    290 GQG---SGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICSAPQSRPTPSPDT 346

Qy    150 ----------EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGC 198
                    |::| ||||:| : | ||||| ::   |||:||| |: : |||||||
Db    347 WPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGC 406

Qy    199 PSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCY--- 255
           : : |         || |||: | | |   ||:| | |   | :|:| |:|| :  ||
Db    407 GWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAH 466

Qy    256 --------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNA 294
                              | | |||| | :|| ||||:  :| |||||   |:
Db    467 SWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTN 526

Qy    295 AADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQ 354
          |:|||:||||| |:   |        ||  : || | ||||||:|| | |:| :: :| |
Db    527 DANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHNGWLSHNCGHS 586

Qy    355 NDVSVICS-DGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCP 413
           | ||||   : | ||| :| : | |||||·    | |:|| :|     | |||:||||
Db    587 EDAGVICSGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCG 646

Qy    414 FSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICS---- 469
          :::      |:     :  | :: : |:|:|| ||| |  :|        ||||||
Db    647 WAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPNNGWLSHNCGHHEDAGVICSAAQS 706

Qy    470 -------------------DKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRN 509
                             :: | ||||    | ||:|| |:| ||||| | | | | :
Db    707 RSTPRPDTLSTITLPPSTVGSESSLTLRLVNGSDRCQGRVEVLYRGSWGTVCDDSWDTND 766

Qy    510 AAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHRE 569
          | |||:|||||     |    | : |||| |||| | |:|| :| | | |:||   ||| | |
Db    767 ANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHE 826

Qy    570 DVIVTCS--------GDATW-------------GLRLVGGSNRCSGRLEVYFQGRWGTV 607
          |  | ||           ||             |||| | :|| ||:|| ::| ||||
Db    827 DAGVICSVSQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTV 886

Qy    608 CDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGW 667
          ||| |::  ||| || |  :     |       | | | |||| | | || |||| ::||
Db    887 CDDSWDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGW 946

Qy    668 GNNDCSHSEDVGVICSDA-----------------------SDMELRLVGGSSRCAGKV 703
          :::| |||| ||||| |                        | : |||| |  || |:|
Db    947 LSHNCQHSEDAGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRV 1006

Qy    704 EVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGE 763
          ||  ||: | :| : |  | | |||||| || |:           | : :  |:: :: |:| |
Db   1007 EVLYQGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDARCSGHE 1066

Qy    764 ASLWDCIRWEWKQTACHLNMEASLICSAHR-QP----------------------RLVG 799
          : ||  |   .  |  | : :| :|||| : :|                      |||
Db   1067 SYLWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVN 1126

Qy    800 ADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWA 859
            | |||||  :  :| :|||   :     |||  |||:| ||  |:|      ||:|:|
Db   1127 GGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVACRQLGCGWAMSAPGNARFGQGSGPIVL 1186

Qy    860 EKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTD------------------- 899
          : :| | |::| ||          | |  : ||:||
Db   1187 DDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSE 1246

Qy    900 ----VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYI 955
              :||||| :| |:||:  | ||::|| :|| || |:|||| || | | |
Db   1247 SSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARF 1306

Qy    956 GERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVS 1015
          |: |   :        | :|| | |   :      | |        |||: | :|| |
Db   1307 GQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPT-PSPDTWP 1365

Qy   1016 DPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQ 1075
           : |     |:|        ||||:| || |||: : | |||:||| || :||:|||:
Db   1366 TSHASTAGSESSL------ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCR 1419

Qy   1076 KLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC 1135
          :|||| | :|  :| ||:|||||| |||: |:| ||:|| ||   ||  |:| | ||||||
Db   1420 QLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVIC 1479

Qy   1136 SEF----------------------TALRLYSETETESCAGRLEVFYNGTWGSVGRRN 1171
          | |                      |||| :    : | ||:|| | |:||:|
Db   1480 SAFQSQPTPSPDTWPTSRASTAGSESTLALRLVN--GGDRCRGRVEVLYQGSWGTVCDDY 1537

Qy   1172 ITTAIAGIVCRQLGCGENGVVSLAPLSK---TGSGFMWVDDIQCPKTHISIWQCLSAPW- 1227
           |  | :|||||||   |   ||   |:     ||| : :||:|: |||  :|| |
Db   1538 WDTNDANVVCRQLGC---GWAMSAPGNAQFGQGSGPIVLDDVRCSGHEPYLWSCPHNGWL 1594

Qy   1228 ----------------ERRISSPAEETWIT---------CEDRIRVR---GGDTECSGRV 1259
                            : |:| :||:|              | : :|    ||| | |||
Db   1595 SHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGD-RCRGRV 1653

Qy   1260 EIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNE 1319
          |: : ||||||||||| :| ||:||||| |:::|  :| ||||:| | | |:|| |||
Db   1654 EVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLGDVRCSGNE 1713

Qy   1320 SFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
          |:|| |  | |  :||| |||| ||    : |       |
Db   1714 SYLWSCPHKGWLTHNCGHHEDAGVICSATQINSTTTDWWH 1753
```

```
RESULT 10
Q96DU4
ID   Q96DU4      PRELIMINARY;      PRT;  2413 AA.
AC   Q96DU4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DMBT1/8kb.2 protein precursor.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMALL INTESTINE;
RA   Mollenhauer J.;
RT   "Major subforms of DMBT1 are gastrointestinal mucins that display
RT   extensive alternative splicing and differential protein targeting.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ297935; CAC44122.1; -.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 14.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PROSITE; PS01180; CUB; 2.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR   PROSITE; PS50287; SRCR_2; 14.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW   Signal.
FT   SIGNAL          1     25       POTENTIAL.
```

```
FT   CHAIN        26   2413       DMBT1/8KB.2 PROTEIN.
SQ   SEQUENCE   2413 AA;  260600 MW;  99A449C5F4F60728 CRC64;

  Query Match            36.6%;  Score 2982;  DB 4;  Length 2413;
  Best Local Similarity  40.1%;  Pred. No. 8.7e-224;
  Matches  618;  Conservative  176;  Mismatches  506;  Indels  240;  Gaps   23;

Qy         44 TDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM---FRF 100
                : | ||||||         | | |||  ::| |||||| |:|   :   |||:||||  :: :     :|
Db        230 SSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQF 289

Qy        101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG----------- 149
              ||       | |  ||||  | |:|| ||  |  |    |  :|||  |  ||  ||  |
Db        290 GQG---SGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICSAPQSRPTPSPDT 346

Qy        150 -----------EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGC 198
                         |::|  ||||:|  :  |||||  ::    |||:|||  |:  :  | |||||||
Db        347 WPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGC 406

Qy        199 PSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCY--- 255
                :   :  |            || |||:  |  |  |    |||:|  |  ||   :|:|  |:||   :  |
Db        407 GWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAH 466

Qy        256 --------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNA 294
                                  |  |  ||||  |  :||  ||||:   :|  |||||     |:
Db        467 SRSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTN 526

Qy        295 AADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQ 354
                |:|||:|||||  |:     |           ||     :  ||  |  |||||||:||  |  |:|  ::  :|  |
Db        527 DANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHNGWLSHNCGHS 586

Qy        355 NDVSVICS-DGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCP 413
                |   ||||   ':  |  |||  :|  :  |  |||||        |  |:||  :|       |  |||:||||
Db        587 EDAGVICSGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCG 646

Qy        414 FSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICS---- 469
              ::           |:      :     |  :  :  |:|:||  ||  |       :|         |  |||||||
Db        647 WATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPNNGWLSHNCGHHEDAGVICSAAQS 706

Qy        470 -------------------DKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRN 509
                                 :: |  ||||         |  ||:||  |:|  |||||  |  |  |  :
Db        707 RSTPGPDTLSTITLPPSTVGSESSLTLRLVNGSDRCQGRVEVLYRGSWGTVCDDSWDTND 766

Qy        510 AAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHRE 569
                | |||:|||||        |       |  :  ||||  ||||  |  |:||  :|  |  |:||    |||  |  |
Db        767 ANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHE 826

Qy        570 DVIVTCS--------GDATW--------------GLRLVGGSNRCSGRLEVYFQGRWGTV 607
                |   |  ||                   ||                              ||||  |   :||  ||:||   ::|  ||||
Db        827 DAGVICSVSQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTV 886

Qy        608 CDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGW 667
              |||  |::       |  |||  ||  |       :                 |     |  |  ||||  |  |  |  ||  ||||   ::||
Db        887 CDDSWDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGW 946

Qy        668 GNNDCSHSEDVGVICSDA------------------------SDMELRLVGGSSRCAGKV 703
                :::|  ||||  |||||  |                                   |  :  ||||  |     ||  |:|
Db        947 LSHNCQHSEDAGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRV 1006

Qy        704 EVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGE 763
              ||   ||:  |  :|  :    |     |  |  |||||||  ||  |:                 |  :  :     |::  :     |:|  |
Db       1007 EVLYQGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHE 1066

Qy        764 ASLWDCIRWEWKQTACHLNMEASLICSAHR-QP-----------------------RLVG 799
              :  ||  |          |       |      :  :|  :||||  :  :|                                      |||
Db       1067 SYLWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVN 1126

Qy        800 ADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWA 859
                  |  |||||   :     :  :||   :    :  |||:|||:|   ||  |:|            '||:|:|
Db       1127 GGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVL 1186

Qy        860 EKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTD-------------------- 899
                :   :|  |   |::|   ||                 |  |    :   ||:||
Db       1187 DDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSE 1246

Qy        900 ----VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYI 955
                  :||||   :|  |:||:     |  ||::||  :||   ||  |:||'|   |    |    |    |
Db       1247 SSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCSWATSAPGNARF 1306

Qy        956 GERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVS 1015
              |:  |     :         |  |:||  |  :|            :      |    |            |||:  |   :||    |
Db       1307 GQGSGPIVLDDVRCSGHESYLWSCPHNGWFSHNCGHHEDAGVICSASQSQPT-PSPDTWP 1365

Qy       1016 DPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQ 1075
                :  |        |:|           ||||:|    ||  ||||:   :   |   |||:|||   ||   :||:|||:
Db       1366 TSHASTAGSESSL------ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCR 1419

Qy       1076 KLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC 1135
              :||||  |  :|    :|  ||:||||||  |||:  |:|  ||:||  ||   ||  |:|  |  ||||||
Db       1420 QLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVIC 1479

Qy       1136 ----------------------SEFT-ALRLYSETETESCAGRLEVFYNGTWGSVGRRN 1171
                                    ||  |  ||||  :        :  |  ||:||  |  |:||:|
Db       1480 SASQSQPTPSPDTWPTSRASTAGSESTLALRLVN--GGDRCRGRVEVLYQGSWGTVCDDY 1537

Qy       1172 ITTAIAGIVCRQLGCGENGVVSLAPLSK---TGSGFMWVDDIQCPKTHISIWQCLSAPW- 1227
                |   |  :|||||||      |         ||  :         |||  :  :||::|               :|  |       |
Db       1538 WDTNDANVVCRQLGC---GWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWL 1594

Qy       1228 ----------------ERRISSPAEETWIT---------CEDRIRVR---GGDTECSGRV 1259
                              :    |:|    :||:|                      |      :   :|         |||     |||
Db       1595 SHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGD-RCRGRV 1653

Qy       1260 EIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNE 1319
              |:  :  ||||||||||||   :|  |||:|||||  |::|   :|·||||:|  |  |||:||  |||
Db       1654 EVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNE 1713

Qy       1320 SFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
              |:||  |      |  |      :|||  |||||  ||          :  |                  |
Db       1714 SYLWSCPHKGWLTHNCGHHEDAGVICSATQINSTTTDWWH 1753

RESULT 11
Q9UKJ4
ID   Q9UKJ4      PRELIMINARY;      PRT;   2413 AA.
AC   Q9UKJ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Gp-340 variant protein.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99415938; PubMed=10485905;
RA   Holmskov U., Mollenhauer J., Madsen J., Vitved L., Gronlund J.,
RA   Tornoe I., Kliem A., Reid K.B., Poustka A., Skjodt K.;
RT   "Cloning of gp-340, a putative opsonin receptor for lung surfactant
RT   protein D.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC   -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR   EMBL; AF159456; AAD49696.1; -.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 14.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00042; CUB; 2.
DR   SMART; SM00202; SR; 14.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS01180; CUB; 2.
```

```
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR   PROSITE; PS50287; SRCR_2; 14.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
SQ   SEQUENCE   2413 AA;  260755 MW;  039544043CF463D4 CRC64;

  Query Match             36.6%;  Score 2980;  DB 4;  Length 2413;
  Best Local Similarity   40.1%;  Pred. No. 1.2e-223;
  Matches  618;  Conservative  177;  Mismatches  505;  Indels  240;  Gaps   23;

Qy     44 TDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM---FRF 100
Db    230 SSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQF 289

Qy    101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG----------- 149
Db    290 GQG---SGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICSAPQSRPTPSPDT 346

Qy    150 -----------EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGC 198
Db    347 WPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGC 406

Qy    199 PSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCY--- 255
Db    407 GWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSAAH 466

Qy    256 --------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNA 294
Db    467 SWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTN 526

Qy    295 AADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQ 354
Db    527 DANVVCRQLGCGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHNGWLSHNCGHS 586

Qy    355 NDVSVICS-DGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCP 413
Db    587 EDAGVICSGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCG 646

Qy    414 FSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICS---- 469
Db    647 WAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPNNGWLSHNCGHHEDAGVICSAAQS 706

Qy    470 -------------------DKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRN 509
Db    707 RSTPRPDTLSTITLPPSTVGSESSLTLRLVNGSDRCQGRVEVLYRGSWGTVCDDSWDTND 766

Qy    510 AAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHRE 569
Db    767 ANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHE 826

Qy    570 DVIVTCS--------GDATW-------------GLRLVGGSNRCSGRLEVYFQGRWGTV 607
Db    827 DAGVICSVSQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTV 886

Qy    608 CDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGW 667
Db    887 CDDSWDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGW 946

Qy    668 GNNDCSHSEDVGVICSDA-----------------------SDMELRLVGGSSRCAGKV 703
Db    947 LSHNCQHSEDAGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRV 1006

Qy    704 EVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGE 763
Db   1007 EVLYQGSWGTVCDDSWDTNDANVVCRQPGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHE 1066

Qy    764 ASLWDCIRWEWKQTACHLNMEASLICSAHR-QP----------------------RLVG 799
Db   1067 SYPWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVN 1126

Qy    800 ADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWA 859
Db   1127 GGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVL 1186

Qy    860 EXFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTD------------------- 899
Db   1187 DDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSE 1246

Qy    900 ----VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYI 955
Db   1247 SSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARF 1306

Qy    956 GERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVS 1015
Db   1307 GQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPT-PSPDTWP 1365

Qy   1016 DPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQ 1075
Db   1366 TSHASTAGSESSL------ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCR 1419

Qy   1076 KLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC 1135
Db   1420 QLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVIC 1479

Qy   1136 ---------------------SEFT-ALRLYSETETESCAGRLEVFYNGTWGSVGRRN 1171
Db   1480 SASQSQPTPSPDTWPTSRASTAGSESTLALRLVN--GGDRCRGRVEVLYQGSWGTVCDDY 1537

Qy   1172 ITTAIAGIVCRQLGCGENGVVSLAPLSK---TGSGFMWVDDIQCPKTHISIWQCLSAPW- 1227
Db   1538 WDTNDANVVCRQLGC---GWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWL 1594

Qy   1228 ----------------ERRISSPAEETWIT---------CEDRIRVR---GGDTECSGRV 1259
Db   1595 SHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGD-RCRGRV 1653

Qy   1260 EIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNE 1319
Db   1654 EVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNE 1713

Qy   1320 SFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
Db   1714 SYLWSCPHKGWLTHNCGHHEDAGVICSATQINSTTTDWWH 1753


RESULT 12
Q9Y211
ID   Q9Y211      PRELIMINARY;      PRT;  1785 AA.
AC   Q9Y211;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DMBT1.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20017478; PubMed=10551316;
RA   Takeshita H., Sato M., Shiwaku H.O., Semba S., Sakurada A., Hoshi M.,
RA   Hayashi Y., Tagawa Y., Ayabe H., Horii A.;
RT   "Expression of the DMBT1 gene is frequently suppressed in human lung
RT   cancer.";
RL   Jpn. J. Cancer Res. 90:903-908(1999).
CC   -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR   EMBL; AB020851; BAA78577.1; -.
DR   EMBL; AB020812; BAA78577.1; JOINED.
DR   EMBL; AB020813; BAA78577.1; JOINED.
DR   EMBL; AB020814; BAA78577.1; JOINED.
DR   EMBL; AB020815; BAA78577.1; JOINED.
DR   EMBL; AB020816; BAA78577.1; JOINED.
DR   EMBL; AB020817; BAA78577.1; JOINED.
DR   EMBL; AB020818; BAA78577.1; JOINED.
DR   EMBL; AB020819; BAA78577.1; JOINED.
DR   EMBL; AB020820; BAA78577.1; JOINED.
```

```
DR   EMBL; AB020821; BAA78577.1; JOINED.
DR   EMBL; AB020822; BAA78577.1; JOINED.
DR   EMBL; AB020823; BAA78577.1; JOINED.
DR   EMBL; AB020824; BAA78577.1; JOINED.
DR   EMBL; AB020825; BAA78577.1; JOINED.
DR   EMBL; AB020826; BAA78577.1; JOINED.
DR   EMBL; AB020827; BAA78577.1; JOINED.
DR   EMBL; AB020828; BAA78577.1; JOINED.
DR   EMBL; AB020829; BAA78577.1; JOINED.
DR   EMBL; AB020830; BAA78577.1; JOINED.
DR   EMBL; AB020831; BAA78577.1; JOINED.
DR   EMBL; AB020832; BAA78577.1; JOINED.
DR   EMBL; AB020833; BAA78577.1; JOINED.
DR   EMBL; AB020834; BAA78577.1; JOINED.
DR   EMBL; AB020835; BAA78577.1; JOINED.
DR   EMBL; AB020836; BAA78577.1; JOINED.
DR   EMBL; AB020837; BAA78577.1; JOINED.
DR   EMBL; AB020838; BAA78577.1; JOINED.
DR   EMBL; AB020839; BAA78577.1; JOINED.
DR   EMBL; AB020840; BAA78577.1; JOINED.
DR   EMBL; AB020841; BAA78577.1; JOINED.
DR   EMBL; AB020842; BAA78577.1; JOINED.
DR   EMBL; AB020843; BAA78577.1; JOINED.
DR   EMBL; AB020844; BAA78577.1; JOINED.
DR   EMBL; AB020845; BAA78577.1; JOINED.
DR   EMBL; AB020846; BAA78577.1; JOINED.
DR   EMBL; AB020847; BAA78577.1; JOINED.
DR   EMBL; AB020848; BAA78577.1; JOINED.
DR   EMBL; AB020849; BAA78577.1; JOINED.
DR   EMBL; AB020850; BAA78577.1; JOINED.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 9.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00042; CUB; 2.
DR   SMART; SM00202; SR; 9.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS01180; CUB; 2.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_8.
DR   PROSITE; PS50287; SRCR_2; 9.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
SQ   SEQUENCE   1785 AA;  193991 MW;  38B2363F95226EB0 CRC64;

  Query Match            30.1%;  Score 2448;  DB 4;  Length 1785;
  Best Local Similarity  36.2%;  Pred. No. 3.2e-182;
  Matches  534;  Conservative  156;  Mismatches  413;  Indels  374;  Gaps   28;

Qy     43 GTD--LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM--- 97
Db     95 GSDSGLALRLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGN 154

Qy     98 FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCY--------- 148
Db    155 AWFGQG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLR 211

Qy    149 ---------------GEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVV 192
Db    212 PESWPVRISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVV 271

Qy    193 CRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTL 252
Db    272 CRQLGCGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGV 331

Qy    253 TCY---------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHK 290
Db    332 ICSAPQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDS 391

Qy    291 WNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFD 350
Db    392 WDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHN 451

Qy    351 CLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQL 410
Db    452 CQHSEDAGVICSAA----------------------HSWST----------------- 470

Qy    411 GCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470
Db    471 ------------PSPDTLPTITLPASTVGSESS------------------------ 491

Qy    471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTY 530
Db    492 ---LALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGNAR 548

Qy    531 FKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDA--------TW- 581
Db    549 FGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWP 608

Qy    582 ------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPS 628
Db    609 TSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGW 668

Qy    629 SIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICS----- 683
Db    669 AMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQ 728

Qy    684 ----------------DASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEV 726
Db    729 PTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANV 788

Qy    727 VCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEAS 786
Db    789 VCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAG 848

Qy    787 LICSAHR-QPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLS 845
Db    849 VICSASQSQP-------------TPSPDTWPT------------------------- 867

Qy    846 VGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNG 905
Db    868 --------------SRASTAGSESTLAL------------------------RLVNG 886

Qy    906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGH 965
Db    887 GDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQFGQGSGPIVLD 946

Qy    966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPYLSAVP 1023
Db    947 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGS 1006

Qy   1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAF 1083
Db   1007 ESSL-------ALRLVNGGDRCRGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAM 1059

Qy   1084 NATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRL 1143
Db   1060 SAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLTHNCGHHEDAGVICS---ATQI 1116

Qy   1144 YSET------------ETESCAGRLEVFY-NGTWGS--------------------- 1166
Db   1117 NSTTTDWWHPTTTTARPSSNCGGFL--FYASGTFSSPSYPAYYPNNAKCVWEIEVNSGY 1174

Qy   1167 ---VGRRNI--------------------TTAIAGIVC---RQLGCGENGVVSL---AP 1196
Db   1175 RINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICNDTRQIFTSSYNRMTIHFRSD 1234

Qy   1197 LSKTGSGFM-WVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTEC 1255
Db   1235 ISFQNTGFLAWYN---------------------SFPSDATL----RLVNLNSSYGLC 1267

Qy   1256 SGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRC 1315
Db   1268 AGRVEIYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVEC 1327
```

```
Qy   1316 KGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352
              | || || |   :  |    :| |:||||| |||   |  :
Db   1328 SGTESTLWQCRNRGWFSHNCNHREDAGVICSGNHLST 1364


RESULT 13
O97375
ID   O97375      PRELIMINARY;      PRT;  2153 AA.
AC   O97375;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Scavenger receptor cysteine-rich protein type 12 precursor.
GN   SRCR12.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99328904; PubMed=10398804;
RA   Pancer Z., Rast J.P., Davidson E.H.;
RT   "Origins of immunity: transcription factors and homologues of effector
RT   genes of the vertebrate immune system expressed in sea urchin
RT   coelomocytes.";
RL   Immunogenetics 49:773-786(1999).
DR   EMBL; AF064259; AAD08654.1; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 20.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 20.
DR   PROSITE; PS00420; SRCR_1; 15.
DR   PROSITE; PS50287; SRCR_2; 20.
KW   Receptor; Signal.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17   2153       SCAVENGER RECEPTOR CYSTEINE-RICH PROTEIN
FT                                TYPE 12.
SQ   SEQUENCE   2153 AA;  226557 MW;  98C51996A272A010 CRC64;

  Query Match             30.0%;  Score 2439.5;  DB 5;  Length 2153;
  Best Local Similarity   38.2%;  Pred. No. 1.9e-181;
  Matches  527;  Conservative  197;  Mismatches  560;  Indels   97;  Gaps   22;

Qy     32 LNSCFLISSFN-------------GTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWN 78
          |:||  :|| |                |   :||| |      | ||:    |||||||| |:
Db    113 LSSCAYVSSHNCAHSEDAGVVCAGSGDGNIRLVGGSHSSEGRVELYVNNQWGTVCDDLWD 172

Qy     79 TTASTVVCKQLGC-PFSFAMFRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNC 137
            :  | |:|||  |      ||:         ||| |:|:| |   |:::|  |   |     ||||
Db    173 LNDAQVACRQLGLGPAVSADAHFGEG---SGSILLDNVACIGTETSLLSCSHPGIESHNC 229

Qy    138 YHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLG 197
           | || || |         :|:| |:      |||||       |||:||| |::| |:| |||||
Db    230 GHSEDAGVICSSRFGSDIRVVGGSIPTEGRVEVFVNGAWGTVCDDLWDINDASVACRQLG 289

Qy    198 CPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDS 257
            :  :||  :          |:||:: | | |  | :| | |  |:|:| |  ||   :  |  |
Db    290 FGRAISASGGASYGQGSGSIFLDNLACTGAESNLLSCPHNGVGSHNCGHGEDAGVLCSSS 349

Qy    258 SDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHL 317
           ::|:||| :   ||||: :    |||||    |:    :| |:|||  | |:   |
Db    350 LSSDIRVVGGRSPTEGRVEVFVNRAWGTVCDDLWDVNDTNVACRQLGFGRAISAPGGASY 409

Qy    318 QSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGA-DLELRLADGSN 376
            ||  : || ::|:| || |  | |:|   :| |   |  | |:   :  :  :|| | |
Db    410 GQGSGSILLDNLACTGAESSLLSCPHNGVGSHNCAHGEDAGVSCAPSSQESRVRLVGGLN 469

Qy    377 NCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKP--SNEARDIWIN 434
           |  ||||: ::  || |:|| :|       | |||:|||  |    || |:              :  | ::
Db    470 NREGRVEIFLNNQWGTVCDDDWGTPDANVVCRQLGYPSG--GSARSSAYFGRGSVPILLD 527

Qy    435 SISCTGNESALWDCTYDGKAKRTCFRRSDAGVIC---SDKADLDLRLVGAHSPCYGRLEV 491
           :: |:|||  :|  |: :|       |  : || |:|      :    :::|:||  ||  ||:||
Db    528 NVGCSGNERSLELCSNNGIGVHNCGHQEDASVVCTGVTTMPGMNIRIVGGGSPSEGRVEV 587

Qy    492 KYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESN 551
                  ||||| | |     :| |||::||              | :| || | |||:  | | ||:
Db    588 LVGHRWGTVCDDLWDINDANVVCRELGYSAATSATSSASFGQGSGDILLDDLRCSGTESS 647

Qy    552 IWDCEHSGWGKHNCVHREDVIVTCSGDATWG------LRLVGGSNRCSGRLEVYFQGRWG 605
          :  | | |  | | ||| | ||  | |:    | |         :||||.|   ||||:    :||
Db    648 LLTCPHRGVGVHNCAHSEDAGVVCA-SITSGPVVGTYMRLVGGQNSRQGRLEISINNQWG 706

Qy    606 TVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNS 665
          ||||| |:   | ||| ||   |::          :    | | |||||||| |:|: |  | :
Db    707 TVCDDSWDINDATVVCRQLGFSSAVSAPTSAHFGQGSGTIWLDDVSCAGNENSLMDCGHR 766

Qy    666 GWGNNDCSHSEDVGVICSDASD--MELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNI 723
           | | ::|:|:|| ||:|  |||  : :|| || |  |    |:||::: |  | :| : ||:
Db    767 GLGVHNCAHAEDAGVVCI-ASDGPLNIRLAGGRSGMEGRVEISLGGDWGTVCDDSWGIED 825

Qy    724 AEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNM 783
          | |||||| | |:          | : :  ||| |   |:| ||: :|   |
Db    826 AHVVCRQLGFGPALSAVTAASFGQGSGSILMDNVQCSGDEATIAECSHNGIGIHNCGHQE 885

Qy    784 EASLICS-AHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
          :| ::|| |    : ||||            ||||: :   | :|||  :  |   | |:|:|||: |::||
Db    886 DAGVVCSRATGEVRLVGNSRQNEGRVEILISGRWGTVCDDLWDLRDADVICRQLGYGNAI 945

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP---EDTCIHSREVGVVCSRYTD 899
          |    || | |       :    | |||   :   |     ||       :  | || : ||||     :
Db    946 SAPHSSFFGPGRGDILLDDVSCTGSEDDILDC---SHPPIGTNNCGHSEDAGVVCD--LN 1000

Qy    900 VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERS 959
          ||:|||    :|:|||  |  | ::|| :|| :|| |:||||  :|         |: |
Db   1001 VRIVNGSRSNEGRVEILHDGSWATVCDDNWDIDDATVVCRQLGFPSANYAVAHAVFGQGS 1060

Qy    960 VRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYL 1019
            :    | |:| |  ||  ||  | |    |||:        ||
Db   1061 GEIVLDDVECTGDEVSLIECQHAGLGTNNCGHSEDAGVICS-----------VNV----- 1104

Qy   1020 SAVPEGSALICLEDKRLRLVGDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGC 1079
                          || ||:|    |||||:| || |||:|||  | : ||:|: :|
Db   1105 -----------------RLADGNSPAEGRVEVYFDGQWGTVCDDNWGLQNGHVICRAVGF 1147

Qy   1080 GVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFT 1139
           | |   | || | |||:||::||||:|: |  |   | | ||| | |||||||||: .
Db   1148 GKAIAVTDRARFGRGQGPIYLDNINCTGSETSLLNCSHSGLGTHDCDHSEDAGVICSQPD 1207

Qy   1140 ALRL---YSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAP 1196
           :||  |:| ||:|: || ||:|  | :|||:||  :
Db   1208 TIRLAGGRSKYE-----GRVEILQNGAWGTVCDNQWDVHDAAVVCRELGYATTLEATSQA 1262

Qy   1197 LSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVR--GGDTE 1254
            |:| : :||::|   | :  |   :  |:  : |:| |  :|  | |
Db   1263 SFGQGTGAILLDDLRCSGREIRLVDCPHNGLGQHNCQHLEDAGVVCQDVIELRLVDGQTP 1322

Qy   1255 CSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMR 1314
          :|||: :  ||||||: || :| |||::|| |:|| |  :| || |:   |: |
Db   1323 NAGRVELRYQGQWGTVCDNGWDVRDAAVVCRELGFPRAIAANNQAFYGPGIGTVLLSDIS 1382

Qy   1315 CKGNESFLWDCHAKPWGQSDCGHKEDAGVRC------SGQSLKSLNASSGHLALILSSIF 1368
           | | :| |: |      | : | | ||||| |         ||   | | : | : :: ::
Db   1383 CGGQEATLFSCPNSGVGLNSCNHGEDAGVVCMPKVRLSG----SNNPNEGLVEVFMNGLW 1438

Qy   1369 G 1369
          |
Db   1439 G 1439


RESULT 14
Q9Y4V9
ID   Q9Y4V9      PRELIMINARY;      PRT;  1785 AA.
AC   Q9Y4V9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DMBT1/6kb.1 protein precursor.
GN   DMBT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE=97434209; PubMed=9288095;
RA   Mollenhauer J., Wiemann S., Scheurlen W., Korn B., Hayashi Y.,
RA   Wilgenbus K.K., von Deimling A., Poustka A.;
RT   "DMBT1, a new member of the SRCR superfamily on chromosome 10q25.3-
RT   q26.1 is deleted in malignant brain tumours.";
RL   Nat. Genet. 17:32-39(1997).
CC   -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR   EMBL; AJ000342; CAA04019.1; -.
DR   InterPro; IPR000859; CUB_domain.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00431; CUB; 2.
DR   Pfam; PF00530; SRCR; 9.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00042; CUB; 2.
DR   SMART; SM00202; SR; 9.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS01180; CUB; 2.
DR   PROSITE; PS00420; SRCR_1; UNKNOWN_8.
DR   PROSITE; PS50287; SRCR_2; 9.
DR   PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW   Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26   1785       DMBT1/6KB.1 PROTEIN.
SQ   SEQUENCE   1785 AA;  193941 MW;  7B1F8D47E4A82092 CRC64;

  Query Match             29.9%;  Score 2435;  DB 4;  Length 1785;
  Best Local Similarity   36.1%;  Pred. No. 3.3e-181;
  Matches  533;  Conservative  155;  Mismatches  415;  Indels  374;  Gaps   28;

Qy      43 GTD--LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM--- 97
           |:|   ||||||||| | | ||: ::| |||||| |:|  : |||:|||| :: :
Db      95 GSDSGLALRLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGN 154

Qy      98 FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCY--------- 148
             |||      | | |||| | |:|| || | |   | |||| |||| || |
Db     155 AWFGQG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLR 211

Qy     149 ---------------GEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVV 192
                          |::| ||||:| : | ||||| ::  |||:||| |: | | ||
Db     212 PESWPVRISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVV 271

Qy     193 CRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTL 252
           ||||||  :  |                 || |||: | |:|  ||:| | || |:| |:|| :
Db     272 CRQLGCGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGV 331

Qy     253 TCY--------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHK 290
            |                      | | |||| | :|| ||||:  :| |||||
Db     332 ICSAPQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDS 391

Qy     291 WNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFD 350
           |: : |:|||:||||| |     |      ||  : || | ||| ||:|| | |:| :: :
Db     392 WDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHN 451

Qy     351 CLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQL 410
           | |  |  ||||                                     ||
Db     452 CQHSEDAGVICSAA----------------------HSWST------------------ 470

Qy     411 GCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470
                         || :          |  |:||:
Db     471 -------------PSPDTLPTITLPASTVGSESS------------------------- 491

Qy     471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTY 530
              | ||||      | ||:|| ||| ||||| | | | :| |||:| |||   |    |
Db     492 ---LALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVVCRQPGCGWAMSAPGNAR 548

Qy     531 FKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDA--------TW- 581
           | : |||| |||| | |:||  | | |:||  ||| | ||  | || |          ||
Db     549 FGQGSGPIVLDDVRCSGHESYPWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWP 608

Qy     582 ------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPS 628
                       |||| | :|| ||:|| ::| ||||||| |::  | ||| || |
Db     609 TSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGW 668

Qy     629 SIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICS----- 683
           ::   |           | | |  |||| | | || |||| ::|| :::| | || |||||
Db     669 AMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQ 728

Qy     684 -----------------DASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEV 726
                            | : |||| |  || |:||| :|: | :| : | | | |
Db     729 PTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANV 788

Qy     727 VCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEAS 786
           ||||| || |           | : :  |:: : |:| |: || |       |   : :|
Db     789 VCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAG 848

Qy     787 LICSAHR-QPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLS 845
           :|||| : ||                   '||| :
Db     849 VICSASQSQP-------------TPSPDTWPT--------------------------- 867

Qy     846 VGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNG 905
                                :   |||: |||                        |||||
Db     868 -------------SRASTAGSESTLAL--------------------------RLVNG 886

Qy     906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGH 965
            :| |:||:   | ||::|| :|| || |:|||| || |:| |    |: |  :
Db     887 GDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQFGQGSGPIVLD 946

Qy     966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPYLSAVP 1023
             | |:|| | :|         :  | |      |||: : :|  |          : | |:
Db     947 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGS 1006

Qy    1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAF 1083
           | |            ||||:|  || ||||: : |||:||| || ::|||| :|:|||::|| |
Db    1007 ESSL-------ALRLVNGGDRCRGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAM 1059

Qy    1084 NATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRL 1143
           :|  :| ||:|||| |||: |:| ||:|| || :|| |:| | ||||||||   | ::
Db    1060 SAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLTHNCGHHEDAGVICS---ATQI 1116

Qy    1144 YSET------------ETESCAGRLEVFY-NGTWGS---------------------- 1166
            | |                  : :| |  | || :||: |
Db    1117 NSTTTDWWHPTTTTTARPSSNCGGFL--FYASGTFSSPSYPAYYPNNAKCVWEIEVNSGY 1174

Qy    1167 ---VGRRNI--------------------TTAIAGIVC---RQLGCGENGVVSL---AP 1196
             :|  |:                      :: : | :|   ||:         :::   :
Db    1175 RINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICNDTRQIFTSSYNRMTIHFRSD 1234

Qy    1197 LSKTGSGFM-WVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTEC 1255
           :|  :||: | :                                 | |:: |      : :  |
Db    1235 ISFQNTGFLAWYN----------------------SFPSDATL----RLVNLNSSYGLC 1267

Qy    1256 SGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRC 1315
           :|||||: |:|||||||||| : ||||||:||||| :|  :| || |:| || :|| |||: |
Db    1268 AGRVEIYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVEC 1327

Qy    1316 KGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352
            | || || |  :·|   :| |:||||| |||  | :
Db    1328 SGTESTLWQCRNRGWFSHNCNHREDAGVICSGNHLST 1364


RESULT 15
O96943
ID   O96943     PRELIMINARY;      PRT;   2043 AA.
```

```
AC   096943;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SRCR domain, membrane form 2.
GN   SRCRM2.
OS   Geodia cydonium (Sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Astrophorida; Geodiidae; Geodia.
OX   NCBI_TaxID=6047;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98369060; PubMed=9701562;
RA   Blumbach B., Pancer Z., Diehl-Seifert B., Steffen R., Muenkner J.,
RA   Mueller I., Mueller W.E.G.;
RT   "The putative sponge aggregation receptor: Isolation and
RT   characterisation of a molecule composed of scavenger receptor
RT   cysteine-rich domains and short consensus repeats.";
RL   J. Cell Sci. 111:2635-2644(1998).
DR   EMBL; Y14953; CAA75175.1; -.
DR   HSSP; P10998; 1VVD.
DR   InterPro; IPR001190; Srcr_receptor.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00530; SRCR; 14.
DR   Pfam; PF00084; sushi; 6.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00032; CCP; 6.
DR   SMART; SM00202; SR; 14.
DR   PROSITE; PS00420; SRCR_1; 8.
DR   PROSITE; PS50287; SRCR_2; 14.
SQ   SEQUENCE   2043 AA;  220896 MW;  493319D0DC62FAAE CRC64;

  Query Match             27.6%;  Score 2246.5;  DB 5;  Length 2043;
  Best Local Similarity   32.4%;  Pred. No. 2.1e-166;
  Matches  514;  Conservative  196;  Mismatches  588;  Indels  287;  Gaps   31;

Qy      7 SWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLV-NGDGPCSGTVEVKF 65
Db    114 AWH-SVGSECSHQNNSGVVCT----------------DDSAVRLVEGGDMPNRGRVEVYH 156

Qy     66 QGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGNESALW 125
Db    157 ADEWGSICDDYFGFGEAVVICHELGFPSAEDSHGGGTFPQGDGLIWMDNLQCTGHESSIA 216

Qy    126 ECQHREWGSHNCYHGEDVGVNCYGEAN-----LGLRLVDGNNSCS-------GRVEVKFQ 173
Db    217 ECTFPGWGEHNCQHREDAGVTCTPATHDPFTYIPIRLNNGTTSSNATHGLQYGRVEVFIN 276

Qy    174 ERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLR---PIWLDDILCQGNELA 230
Db    277 NTWGTVCDDGWGIEDATIACRQLGFYGALLASSRSEYAQVAPSSVPIYWDEVLCHGDEVA 336

Qy    231 LWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE-----LRLVGGTNRCM--------GRVEL 277
Db    337 LSECHHRTTGVHDCSHYEDAGVFC-TTNDIQPAAPKVRLVRNNNGSVDEDGHVFEGRVEV 395

Qy    278 KIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESF 337
Db    396 NFYGVWGTICDDYWSLADGDVICKMLGFTYADEVGYSGAFGIGTGPIWMDDVQCDGTETR 455

Qy    338 LWDCRHSGTVNFDCLHQNDVSVICS---DGADLELRLADGSNNCSGRVEVRIHEQWWTIC 394
Db    456 IDHCDFPGWGVHNCYHFEDASVVCSYRQPGSD--VRLVGGSSSMEGRVEVLHNGVWGTVC 513

Qy    395 DQNWKNEQALVVCKQLGCPFSVFGSRRAKP------SNEARDIWINSISCTGNESALWDC 448
Db    514 DDSWGAEDAQVICHQLG----FYGTATAVRGTQFGFASSSQPIWLDDVGCTGSEMYLSDC 569

Qy    449 TYDGKAKRTCFRRSDAGVIC---------------------------------------- 468
Db    570 SSSGFGNHNCYHYEDAGVVCQGTSFAIPVRLSNGTRPNEGTVEIFYQNRWGTICDIFWTL 629

Qy    469 ------------------------------------------------------------ 468
Db    630 YDANVVCRQLGYDGAYNATHSSYYGDSGLTTVTAYTDCFGQESQLANCTGFRYPPYIPSW 689

Qy    469 -------------SDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVC 514
Db    690 YCGDNTVAGVMCIGTEDGPAPIRLVGGSEASEGRVEILVQGQWGTICDDIWDQADAEVVC 749

Qy    515 KQLG-CGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIV 573
Db    750 RQLGYLSEGATALQFAHFGEGSGPIVLDDVNCTGLELYVTDCPSSGFYIHNCAHSEDAGV 809

Qy    574 TCSG-----DATWGLRLVGGS---NRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLD 625
Db    810 RCNARGPQTQYTYPIRLVRNGTVITQNEGTVEILHNGTWSAVCDDYWGYTEAVVACHML- 868

Qy    626 CPSSIIGMGLGNASTGY---------GKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSE 676
Db    869 --------GFATAVRAYTRSFHGAVDGDTFLDNVRCAGTEREIFDCYHSSYISRNCDRTQ 920

Qy    677 DVGVICSD--ASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECG 734
Db    921 QAGVACTNYTASEYPIRLVGGSGPHEGRVEIYYQGVWGTVCDDSWGQPDADVVCRQLGYA 980

Qy    735 SAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQ 794
Db    981 NASRATVRAEFGRGTGEIWLDNVACTGFENSLDECRSNGWGDHNCGHREDAGAVCQGELL 1040

Qy    795 P-RLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKG 853
Db   1041 PIRLRDGGSNLEGRVEVYFNKTWGTVCDDFWDLGDATVVCRQLGYPEAVRAEPFARFGSG 1100

Qy    854 NGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVC------SRYTDVRLVNGKS 907
Db   1101 EGPIWLDDVACVGTETSLFECSSSGLGQHNCQHFEDAGVECNDHITPSPVTDLRLAGGTN 1160

Qy    908 QCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYI--GERSVRVWGH 965
Db   1161 ATEGRVEIFFNNTWGTICDDSWDIRDAEVVCRYLGFESAIEALSNGYFGAGDPDQPIWLD 1220

Qy    966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEG 1025
Db   1221 DVDCFGSETIITSCLTSALGEHNCAHYEDAGVRC--YISHPL------------------ 1260

Qy   1026 SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNA 1085
Db   1261 ----------LRLVNGNTEGEGRVEVFHRGSWGTVCDDHWSEIDANIVCKELGFARAISA 1310

Qy   1086 TVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEF----TAL 1141
Db   1311 SGFATFGEGSGTIWLDDVQCDGTEATIFNCPSSGWGNHNCLHSEDASAICTAIPVNPNPV 1370

Qy   1142 RLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGC-GENGVVSLAPLSKT 1200
Db   1371 QLIDGPSVHE--GRLQIYYNNEWGTVCDDQWGYDEANVVCKSLGFPGADPDNSLLNSYGP 1428

Qy   1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDR---IRVRGGDTECSG 1257
Db   1429 GSGHIWLDDVDCRGDEFFIQDCNHADIGENNCGHYEDVGLRCLPNTLDVRLMNGGNTSAG 1488

Qy   1258 RVEIWHAG-SWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCK 1316
Db   1489 RVEVNYNGEEWGTVCDDNWDIKDADVVCRMLDFKSAQSAPREAFFGDGLGVIWFDDFLCS 1548

Qy   1317 GNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLF 1376
Db   1549 GYEDSLLECSHAGVKVHNCRHSEDASVVCS------------------------------ 1578

Qy   1377 ILFLTWCRVQKQKHLPLRVSTRRRG 1401
Db   1579 --MLTVCRDFSQEHGTFQISNRQFG 1601

Search completed: May 12, 2003, 06:42:56
```

Job time : 71 secs

```
OM nucleic - nucleic search, using sw model

Run on:          May 12, 2003, 02:36:30 ; Search time 7768 Seconds
                                           (without alignments)
                                           17338.798 Million cell updates/sec

Title:           US-09-759-130B-379
Perfect score:   4628
Sequence:        1 gcggccgctcgcgatctaga..........aaacggacgcgtcggtcgac 4628

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:    4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                               SUMMARIES
                      %
Result              Query
  No.     Score     Match Length DB ID                Description
-------------------------------------------------------------------------
    1    4543.4     98.2    4579  9  AF264014          AF264014 Homo sapi
    2    3459.8     74.8    3670  6  AX133829          AX133829 Sequence
    3    2244.2     48.5    2249  9  AK058038          AK058038 Homo sapi
    4    1450.6     31.3    3786  9  HSM130AC1         Z22969 H.sapiens m
    5    1448.4     31.3    4950  9  HSM130AC2         Z22970 H.sapiens m
    6    1448.2     31.3    3703  9  HSM130A           Z22968 H.sapiens m
    7    1409.4     30.5    3707  6  AX281768          AX281768 Sequence
    8    1339.2     28.9    3802  9  HSM130AE          Z22971 H.sapiens m
    9    1270.4     27.5    4405 10  AF274883          AF274883 Mus muscu
   10     886.8     19.2    4475  4  BBWC11MR          X63723 B.bovis WC1
   11     649.8     14.0    2305  4  AF400115          AF400115 Ovis arie
   12     551.8     11.9    1665  4  SSSRP3            X99334 S.scrofa mR
c  13     517.8     11.2  202252  2  AC131207          AC131207 Homo sapi
   14     476.4     10.3    2165  4  AF400122          AF400122 Ovis arie
   15     460.6     10.0    1316  4  SSSRP2            X99333 S.scrofa mR
   16     424.4      9.2    7006  3  AF064259          AF064259 Strongylo
   17     385.8      8.3    1519  4  SSSRP4            X99335 S.scrofa mR
   18     362.2      7.8  177768  2  AC131206          AC131206 Homo sapi
   19       339      7.3     630  4  S76311S1          S76311 Ovis aries
   20     336.8      7.3    3431  3  AF228826          AF228826 Strongylo
   21     333.8      7.2    2423  4  SSSRP5            X99336 S.scrofa mR
   22     330.4      7.1  202252  2  AC131207          AC131207 Homo sapi
   23     318.2      6.9    7656  9  HSA243224         AJ243224 Homo sapi
   24     288.2      6.2    7686  9  AF159456          AF159456 Homo sapi
   25     283.8      6.1    7667  9  HSA297935         AJ297935 Homo sapi
   26     281.8      6.1    7683  9  HSA243212         AJ243212 Homo sapi
   27       268      5.8    6323  3  GCY14953          Y14953 Geodia cydo
   28     246.2      5.3    3187  4  SSSRP1            X99332 S.scrofa mR
   29     245.8      5.3    1654  5  AB051832          AB051832 Gallus ga
   30     224.8      4.9     569  9  HSA118397         Y18397 Homo sapien
   31     199.2      4.3     461  6  AX237293          AX237293 Sequence
   32     197.8      4.3    5802  6  A90822            A90822 Sequence 1
   33     197.8      4.3    5802  6  AR193170          AR193170 Sequence
   34     197.8      4.3    5802  9  HSDMBT1           AJ000342 Homo sapi
   35     197.6      4.3    3344  9  HSNEUROTR         AJ001531 Homo sapi
   36     197.6      4.3    3350  6  A83630            A83630 Sequence 1
   37     197.4      4.3    5108  4  AF043112          AF043112 Oryctolag
   38       195      4.2  205331  2  AC124470          AC124470 Mus muscu
   39     194.8      4.2    6074  3  AF076513          AF076513 Strongylo
   40     193.2      4.2    2525 10  BC031429          BC031429 Mus muscu
   41     193.2      4.2    2614 10  D89871            D89871 Mus musculu
   42     192.8      4.2    1796  3  AF228823          AF228823 Strongylo
   43     191.6      4.1    2356  6  A83632            A83632 Sequence 3
   44     191.6      4.1    2356 10  MMY13192          Y13192 Mus musculu
   45     190.8      4.1     538  4  AF400118          AF400118 Ovis arie

                               ALIGNMENTS

RESULT 1
AF264014
LOCUS       AF264014               4579 bp    mRNA    linear   PRI 28-NOV-2000
DEFINITION  Homo sapiens scavenger receptor cysteine-rich type 1 protein M160
            precursor, mRNA, complete cds, alternatively spliced.
ACCESSION   AF264014
VERSION     AF264014.1  GI:9652086
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4579)
  AUTHORS   Gronlund,J., Vitved,L., Lausen,M., Skjodt,K. and Holmskov,U.
  TITLE     Cloning of a novel scavenger receptor cysteine-rich type I
```

```
            transmembrane molecule (M160) expressed by human macrophages
  JOURNAL   J. Immunol. 165 (11), 6406-6415 (2000)
  MEDLINE   20540099
   PUBMED   11086079
REFERENCE   2  (bases 1 to 4579)
  AUTHORS   Groenlund,J., Vitved,L., Lausen,M., Skjodt,K. and Holmskov,U.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2000) Department of Immunology & Microbiology,
            Institute of Medical Biology, University of Southern Denmark,
            Odense University, Winsloewparken 21.1, Odense 5000, Denmark
FEATURES             Location/Qualifiers
     source          1..4579
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q13.3"
     5'UTR           1..25
     CDS             26..4387
                     /codon_start=1
                     /product="scavenger receptor cysteine-rich type 1 protein
                     M160 precursor"
                     /protein_id="AAF91396.1"
                     /db_xref="GI:9652087"
                     /translation="MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGT
                     DLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQ
                     AVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG
                     NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIW
                     LDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVE
                     LKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGN
                     ESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWT
                     ICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTY
                     DGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTR
                     NAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCV
                     HREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQL
                     DCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVIC
                     SDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVS
                     REPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLV
                     GADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGL
                     TWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVE
                     INVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN
                     ESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALIC
                     LEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVS
                     AHFGEGSGPIWLDDLNCTGMESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSE
                     TETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGF
                     MWVDDIQCPKTHISIWQCLSAPWERRISSTAEETWITCEDRIRVRGGDTECSGRVEIW
                     HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES
                     FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGRLALILSSIFGLLLPVLFIL
                     FLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCED
                     ASDTSLLGVLPASEATK"
     sig_peptide     26..151
     misc_feature    152..469
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 1"
     misc_feature    470..790
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 2"
     misc_feature    791..1111
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 3"
     misc_feature    1112..1432
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 4"
     misc_feature    1433..1753
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 5"
     misc_feature    1754..2074
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 6"
     misc_feature    2075..2395
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 7"
     misc_feature    2396..2710
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 8"
     misc_feature    2711..3025
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 9"
     misc_feature    3026..3118
                     /note="Region: interdomain segment"
     misc_feature    3119..3433
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 10"
     misc_feature    3434..3754
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 11"
     misc_feature    3755..4063
                     /note="Region: scavenger receptor cysteine-rich domain
                     SRCR 12"
     misc_feature    4064..4093
                     /note="Region: connecting peptide"
     misc_feature    4094..4171
                     /note="Region: transmembrane domain"
     misc_feature    4172..4384
                     /note="Region: cytoplasmatic domain"
     3'UTR           4388..4579
     polyA_signal    4456..4461
     polyA_signal    4550..4555
BASE COUNT     1147 a    919 c   1343 g   1170 t
ORIGIN

  Query Match             98.2%;  Score 4543.4;  DB 9;  Length 4579;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 4550;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         21 ACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCA 80
              ||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ACCCATAATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCA 78

Qy         81 TCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCAT 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 TCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCAT 138

Qy        141 CAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTC 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTC 198

Qy        201 TGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAA 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAA 258

Qy        261 CACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTT 320
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        259 CACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCCTTCGCCATGTT 318

Qy        321 TCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTA 378

Qy        381 TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTA 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTA 438

Qy        441 TCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGT 500
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        439 TCATGGAGAAGATGTTGGTGTGAACTGTTACGGTGAAGCCAATCTGGGTTTGAGGCTAGT 498

Qy        501 GGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGAC 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        499 GGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGAAC 558

Qy        561 TATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATG 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 TATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATG 618

Qy        621 TCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTG 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    619 TCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTG 678

Qy    681 GCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGG 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    679 GCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGG 738

Qy    741 ATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAG 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    739 ATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAG 798

Qy    801 TGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAA 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 TGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAA 858

Qy    861 AATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGT 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 AATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGT 918

Qy    921 CGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCA 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 CGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCA 978

Qy    981 GTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCT 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 GTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCT 1038

Qy   1041 TTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTC 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 TTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTC 1098

Qy   1101 TGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTG 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 TGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTG 1158

Qy   1161 TTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTG 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTG 1218

Qy   1221 GAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGG 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 GAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGG 1278

Qy   1281 CAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTG 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 CAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTG 1338

Qy   1341 CACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATG 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATG 1398

Qy   1401 CTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCT 1460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 CTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCT 1458

Qy   1461 TGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGG 1520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 TGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGG 1518

Qy   1521 GACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGG 1580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 GACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGG 1578

Qy   1581 ATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTAT 1640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 ATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTAT 1638

Qy   1641 TTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAG 1700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 TTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAG 1698

Qy   1701 TGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGA 1760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 TGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGA 1758

Qy   1761 TGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGT 1820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGT 1818

Qy   1821 GTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGC 1880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 GTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGC 1878

Qy   1881 TGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGC 1940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 TGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGC 1938

Qy   1941 TTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGA 2000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 TTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGA 1998

Qy   2001 TCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGT 2060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 TCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGT 2058

Qy   2061 TGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAG 2120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 TGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAG 2118

Qy   2121 GTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGG 2180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 GTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGG 2178

Qy   2181 CTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAG 2240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2179 CTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAG 2238

Qy   2241 GGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGG 2300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2239 GGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGG 2298

Qy   2301 CTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGC 2360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2299 CTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGC 2358

Qy   2361 GTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGT 2420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2359 GTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGT 2418

Qy   2421 TGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTC 2480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2419 TGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTC 2478

Qy   2481 TGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTG 2540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   2479 TGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTG 2538

Qy   2541 TGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTG 2600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2539 TGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTG 2598

Qy   2601 GGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCA 2660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2599 GGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCA 2658

Qy   2661 ACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATAC 2720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2659 ACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATAC 2718

Qy   2721 AGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 2780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2719 AGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 2778

Qy   2781 TGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATG 2840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2779 TGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATG 2838
```

```
Qy   2841 CAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAG 2900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2839 CAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAG 2898

Qy   2901 AAGTGTTCGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAA 2960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2899 AAGTGTTCGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAA 2958

Qy   2961 CTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGAT 3020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2959 CTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGAT 3018

Qy   3021 CTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATA 3080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3019 CTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATA 3078

Qy   3081 TTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCT 3140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3079 TTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCT 3138

Qy   3141 AGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGG 3200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3139 AGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGG 3198

Qy   3201 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGG 3260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3199 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGG 3258

Qy   3261 CTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCAT 3320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3259 CTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCAT 3318

Qy   3321 CTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCG 3380
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3319 CTGGCTGGATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCG 3378

Qy   3381 CGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATT 3440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3379 CGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATT 3438

Qy   3441 CACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT 3500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3439 CACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT 3498

Qy   3501 CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGG 3560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3499 CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGG 3558

Qy   3561 CATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCGCTTTATC 3620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3559 CATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCGCTTTATC 3618

Qy   3621 TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTC 3680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3619 TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTC 3678

Qy   3681 CATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGAC 3740
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   3679 CATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCACAGCAGAAGAGAC 3738

Qy   3741 CTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAG 3800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3739 CTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAG 3798

Qy   3801 AGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGC 3860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3799 AGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGC 3858

Qy   3861 CGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGA 3920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3859 CGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGA 3918

Qy   3921 CGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAA 3980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3919 CGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAA 3978

Qy   3981 TGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAA 4040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3979 TGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAA 4038

Qy   4041 GGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4039 GGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4098

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
          || ||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   4099 TCGTTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCCGGTTCTGTTTATTCT 4158

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4159 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4218

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4219 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4278

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAG 4340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4279 GGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAG 4338

Qy   4341 CGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTC 4400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4339 CGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTC 4398

Qy   4401 CAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAAT 4460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4399 CAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAAT 4458

Qy   4461 AAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAG 4520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4459 AAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAG 4518

Qy   4521 GAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCACTGT 4580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4519 GAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCACTGT 4578

Qy   4581 T 4581
          |
Db   4579 T 4579
```

```
RESULT 2
AX133829
LOCUS       AX133829                3670 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 15 from Patent WO0119856.
ACCESSION   AX133829
VERSION     AX133829.1  GI:14139781
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3670)
  AUTHORS   Shimkets,R.A., Fernandes,E., Herrmann,J.L., Liu,X., Yang,M. and
            Boldog,F.L.
  TITLE     Secreted human proteins, polynucleotides encoding them and methods
            of using the same
  JOURNAL   Patent: WO 0119856-A 15 22-MAR-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..3670
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             113..3475
                     /note="unnamed protein product"
                     /codon_start=1
```

```
                  /protein_id="CAC38965.1"
                  /db_xref="GI:14139782"
                  /translation="MKNDADVKIRLENRPNSENRCVWPTHRKMDGADLELRLADGSNN
                  CSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINS
                  ISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKY
                  QGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESN
                  IWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCD
                  DGWNGKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGW
                  GNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAE
                  VVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNM
                  EASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDA
                  ISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTD
                  VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGE
                  RSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVS
                  DPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVV
                  CQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGMESHLWQCPSRGWGQHDCRHKEDA
                  GVICSEFTALRLYSETETGSCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGEN
                  GVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIR
                  VRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGT
                  GTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALI
                  LSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPH
                  GTRTSDDTPNHGCEDASDTSLLGVLPASEATK"
BASE COUNT      947 a    752 c   1073 g    898 t
ORIGIN

  Query Match            74.8%;  Score 3459.8;  DB 6;  Length 3670;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 3464;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1114 GATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTA 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        200 GATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGAAGAGTA 259

Qy       1174 GAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAA 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAA 319

Qy       1234 GCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCT 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCT 379

Qy       1294 AAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAG 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 AAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAG 439

Qy       1354 TCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCA 1413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 TCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCA 499

Qy       1414 GATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCAT 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCAT 559

Qy       1474 AGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCAT 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 AGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCAT 619

Qy       1534 GACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCT 1593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 GACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCT 679

Qy       1594 ATGCATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGGATGAC 1653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 ATGCATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGGATGAC 739

Qy       1654 GTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAG 1713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAG 799

Qy       1714 CATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGC 1773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 CATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGC 859

Qy       1774 CTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGA 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 CTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGA 919

Qy       1834 CGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGC 1893
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        920 CGGTGGGGCACAGTGTGTGATGACGGCTGGAACGGTAAAGCTGCAGCTGTGGTGTGTAGC 979

Qy       1894 CAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTAGAGGATAT 1953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 CAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATAT 1039

Qy       1954 GGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGC 2013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 GGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGC 1099

Qy       2014 AGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGT 2073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 AGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGT 1159

Qy       2074 TCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAA 2133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 TCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAA 1219

Qy       2134 GTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAAC 2193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 GTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAAC 1279

Qy       2194 ATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAG 2253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 ATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAG 1339

Qy       2254 CCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGG 2313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1340 CCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGG 1399

Qy       2314 GAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAAT 2373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1400 GAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAAT 1459

Qy       2374 ATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATG 2433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1460 ATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATG 1519

Qy       2434 CCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCT 2493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1520 CCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCT 1579

Qy       2494 GATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATA 2553
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1580 GATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTGTGGAGATGCCATA 1639

Qy       2554 TCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTC 2613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1640 TCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTC 1699

Qy       2614 CAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGAC 2673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1700 CAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGAC 1759

Qy       2674 ACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTT 2733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1760 ACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTT 1819

Qy       2734 GTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGC 2793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1820 GTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGC 1879

Qy       2794 TCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGC 2853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1880 TCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGC 1939

Qy       2854 TGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTG 2913
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1940 TGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTG 1999

Qy       2914 TGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACA 2973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2000 TGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACA 2059

Qy       2974 GTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGC 3033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2060 GTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGC 2119

Qy       3034 CTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTT 3093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2120 CTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTT 2179

Qy       3094 CCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGAC 3153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2180 CCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGAC 2239

Qy       3154 AGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGAT 3213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2240 AGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGAT 2299

Qy       3214 GACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCC 3273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2300 GACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCC 2359

Qy       3274 TTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGAC 3333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2360 TTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGAC 2419

Qy       3334 CTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAG 3393
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       2420 CTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAG 2479

Qy       3394 CACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGG 3453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2480 CACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGG 2539

Qy       3454 CTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGG 3513
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       2540 CTCTACAGTGAAACTGAAACAGGGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGG 2599

Qy       3514 ACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGG 3573
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2600 ACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGG 2659

Qy       3574 CAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCT 3633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2660 CAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCT 2719

Qy       3634 GGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGC 3693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2720 GGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGC 2779

Qy       3694 CTGTCTGCCCCATGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGT 3753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2780 CTGTCTGCCCCATGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGT 2839

Qy       3754 GAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGG 3813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2840 GAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGG 2899

Qy       3814 CACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTG 3873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2900 CACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTG 2959

Qy       3874 GTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGC 3933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2960 GTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGC 3019

Qy       3934 CAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTA 3993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3020 CAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTA 3079

Qy       3994 TGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGC 4053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3080 TGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGC 3139

Qy       4054 GTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTT 4113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3140 GTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTT 3199

Qy       4114 ATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGG 4173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3200 ATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGG 3259

Qy       4174 TGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCT 4233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3260 TGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCT 3319

Qy       4234 CTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGG 4293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3320 CTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGG 3379

Qy       4294 ACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTG 4353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3380 ACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTG 3439

Qy       4354 TTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAG 4413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3440 TTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAG 3499

Qy       4414 ATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCA 4473
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3500 CTCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCA 3559

Qy       4474 AGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAA 4533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3560 AGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAA 3619

Qy       4534 TTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCACTGTTAAA 4584
              ||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       3620 TTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCCCTGTTAAA 3670
```

```
RESULT 3
AK058038
LOCUS       AK058038                2249 bp    mRNA    linear   PRI 31-OCT-2001
DEFINITION  Homo sapiens cDNA FLJ25309 fis, clone SYN00985, highly similar to
            Homo sapiens scavenger receptor cysteine-rich type 1 protein M160
            precursor.
ACCESSION   AK058038
VERSION     AK058038.1  GI:16554040
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens synovial membrane (knee) cDNA to mRNA, clone_lib:SYN
            clone:SYN00985.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
            Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
            Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
            Yamashita,H., Chiba,Y., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S.,
            Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
            Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
            Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2249)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
```

```
            (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
            full insert sequencing: RAB and Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1..2249
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="SYN00985"
                     /tissue_type="synovial membrane (knee)"
                     /clone_lib="SYN"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      559 a    512 c    638 g    540 t
ORIGIN

  Query Match            48.5%;  Score 2244.2;  DB 9;  Length 2249;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 2246;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2333 GTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCT 2392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCT 60

Qy       2393 GCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTG 2452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTG 120

Qy       2453 AAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTG 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTG 180

Qy       2513 CCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATC 2572
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        181 CCAATGTGCTGTGCAGAGAATTAAACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATC 240

Qy       2573 ACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAA 2632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAA 300

Qy       2633 CTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAG 2692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAG 360

Qy       2693 AAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGT 2752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGT 420

Qy       2753 GTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACT 2812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACT 480

Qy       2813 GGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAA 2872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAA 540

Qy       2873 CCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATT 2932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATT 600

Qy       2933 GCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCT 2992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCT 660

Qy       2993 GTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTC 3052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTC 720

Qy       3053 CATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGA 3112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGA 780

Qy       3113 TCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAG 3172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAG 840

Qy       3173 TAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCG 3232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCG 900

Qy       3233 ATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTG 3292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTG 960

Qy       3293 CTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGG 3352
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        961 CTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAATGG 1020

Qy       3353 AGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGG 3412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGG 1080

Qy       3413 AGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAA 3472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAA 1140

Qy       3473 CAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCA 3532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCA 1200

Qy       3533 GGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGA 3592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGA 1260

Qy       3593 ATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATG 3652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATG 1320

Qy       3653 ACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGC 3712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGC 1380

Qy       3713 GAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGC 3772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGC 1440

Qy       3773 GTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCA 3832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCA 1500

Qy       3833 CAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCT 3892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCT 1560

Qy       3893 GTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCT 3952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCT 1620

Qy       3953 GGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAAC 4012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAAC 1680

Qy       4013 CCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGT 4072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGT 1740

Qy       4073 CGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTG 4132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTG 1800

Qy       4133 GGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCACAAACAAA 4192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1801 GGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAA 1860

Qy    4193 AACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCC 4252
Db    1861 AACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGCGGAGAATTTATTCC 1920

Qy    4253 ATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACA 4312
Db    1921 ATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACA 1980

Qy    4313 CCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCT 4372
Db    1981 CCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCT 2040

Qy    4373 CTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 4432
Db    2041 CTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 2100

Qy    4433 TGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACT 4492
Db    2101 TGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACT 2160

Qy    4493 TGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 4552
Db    2161 TGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 2220

Qy    4553 ATAAAAATTGAATTATGTATATCACTGTT 4581
Db    2221 ATAAAAATTGAATTATGTATATCACTGTT 2249


RESULT 4
HSM130AC1
LOCUS       HSM130AC1            3786 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen cytoplasmic variant 1.
ACCESSION   Z22969
VERSION     Z22969.1  GI:312143
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3786)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE   2  (bases 1 to 3786)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
FEATURES             Location/Qualifiers
     source          1..3786
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3557
                     /codon_start=1
                     /product="M130 antigen cytoplasmic variant 1"
                     /protein_id="CAA80542.1"
                     /db_xref="GI:312144"
                     /db_xref="SPTREMBL:Q07899"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSRYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVA
                     LSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQS
                     QTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICD
                     DSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGW
                     GQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVG
                     VVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSE
                     ETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALK
                     AFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDISVQK
                     TPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQ
                     IQYREMNSCLNADDLDLMNSSENSHESADFSAAELISVSKFLPISGMEKEAILSHTEK
                     ENGNL"
     sig_peptide     102..221
     mat_peptide     222..3554
                     /product="M130 antigen cytoplasmic variant 1"
                     /evidence=experimental
     polyA_signal    3760..3765
     polyA_signal    3764..3769
     polyA_signal    3768..3773
     polyA_site      3783
BASE COUNT     1006 a    746 c   1074 g    960 t
ORIGIN

  Query Match             31.3%;  Score 1450.6;  DB 9;  Length 3786;
  Best Local Similarity   65.4%;  Pred. No. 0;
  Matches 2182;  Conservative    0;  Mismatches 1129;  Indels   24;  Gaps    3;

Qy    1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db     107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy    1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db     167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy    1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
Db     227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy    1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
Db     287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy    1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
Db     347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy    1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
Db     407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy    1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
Db     467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy    1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
Db     527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy    1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538
Db     587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy    1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db     647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy    1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db     707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766
```

```
Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1846

Qy   2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db   1847 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1906

Qy   2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db   1907 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966

Qy   2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db   1967 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 2026

Qy   2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db   2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 2086

Qy   2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db   2087 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 2146

Qy   3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
Db   2147 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 2206

Qy   3099 GGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
Db   2207 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 2266

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db   2267 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 2326

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   2327 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 2386

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   2387 TGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 2446

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGA 3398
Db   2447 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAA 2506

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   2507 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 2566

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   2567 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTACAATGGAGCTTG 2626

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   2627 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 2686

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   2687 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 2746

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   2747 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 2806

Qy   3699 TGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   2807 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 2866

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db   2867 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 2926

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
```

```
Db   2927 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 2986

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938

Db   2987 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 3046

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998

Db   3047 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 3106

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058

Db   3107 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 3166

Qy   4059 GTGCTCTGGACAGTCGCTGAA-----------------ATCACTGAATGCCTCCTCAGG 4100

Db   3167 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 3226

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160

Db   3227 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 3286

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220

Db   3287 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAG 3343

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280

Db   3344 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 3403

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGACACCC 4315

Db   3404 TGATCTGGACCTAATGAATTCCTCAGAAAATTCCC 3438


RESULT 5
HSM130AC2
LOCUS       HSM130AC2               4950 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
ACCESSION   Z22970
VERSION     Z22970.1  GI:312145
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4950)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE   2  (bases 1 to 4950)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
FEATURES             Location/Qualifiers
     source          1..4950
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3572
                     /codon_start=1
                     /product="M130 antigen cytoplasmic variant 2"
                     /protein_id="CAA80543.1"
                     /db_xref="GI:312146"
                     /db_xref="SPTREMBL:Q07900"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSRYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVA
                     LSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQS
                     QTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICD
                     DSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGW
                     GQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVG
                     VVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSE
                     ETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALK
                     AFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDISVQK
                     TPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQ
                     IQYREMNSCLNADDLDLMNSSGLWVLGGSIAQGFRSVAAVEAQTFYFDKQLKKSKNVI
                     GSLDAYNGQE"
     sig_peptide     102..221
     mat_peptide     222..3569
                     /product="M130 antigen cytoplasmic variant 2"
                     /evidence=experimental
     polyA_signal    4924..4929
     polyA_signal    4928..4933
     polyA_signal    4932..4937
     polyA_site      4947
BASE COUNT     1348 a    970 c   1277 g   1355 t
ORIGIN

  Query Match             31.3%;  Score 1448.4;  DB 9;  Length 4950;
  Best Local Similarity   65.4%;  Pred. No. 0;
  Matches 2178;  Conservative    0;  Mismatches 1126;  Indels   24;  Gaps    3;

Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061

Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121

Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181

Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy   1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241

Db    287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy   1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301

Db    347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy   1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361

Db    407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy   1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418

Db    467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy   1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478

Db    527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy   1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538

Db    587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
```

```
Db    647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy   1599 TGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db    707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1846

Qy   2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db   1847 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1906

Qy   2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db   1907 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966

Qy   2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db   1967 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 2026

Qy   2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db   2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 2086

Qy   2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db   2087 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 2146

Qy   3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
Db   2147 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 2206

Qy   3099 GGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
Db   2207 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 2266

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db   2267 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 2326

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   2327 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 2386

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   2387 TGCCACTGGTTCTGCTCATTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 2446

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGA 3398
Db   2447 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAA 2506

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   2507 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 2566

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   2567 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTACAATGGAGCTTG 2626

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   2627 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 2686

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   2687 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 2746

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   2747 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 2806

Qy   3699 TGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   2807 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 2866
```

```
Qy     3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
              | |||||| | |    | ||||  |||      || |||||  | |||||||||||||| |
Db     2867 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 2926

Qy     3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
            ||| |||||||| |||||||||||||||| |||||| ||| |||.||    | ||||||||
Db     2927 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 2986

Qy     3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
            ||| || || |||||||  | ||| ||    ||   |  |  |.| ||    ||||| |||||
Db     2987 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 3046

Qy     3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
            |||||| | || ||| |  ||||   ||  ||||||||| |||||||| |.    | |||||
Db     3047 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 3106

Qy     3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
             ||||  |||| || ||||||  || |||||  ||||| |||||||||.|||| ||||  ||||
Db     3107 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 3166

Qy     4059 GTGCTCTGGACAGTCGCTGAA------------------ATCACTGAATGCCTCCTCAGG 4100
             ||| | |      || || |                   ||.  .|. ||| || |
Db     3167 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 3226

Qy     4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
            ||| | | |   |||||   |  |.| ||| ||||| |  ||||. |||   .| || |
Db     3227 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 3286

Qy     4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
            |||  ||    | |     |    |   | | ||| | |||    ||  ||  . ||||||  | ||
Db     3287 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAG 3343

Qy     4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
             | |    | . |   |   | |  |  |  |||  ||  ||||||  |   | ||||.|  ||   |||
Db     3344 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 3403

Qy     4281 GGACCCACATGGGACAAGAACCTCAGAT 4308
             || |    |      |  |    |||||| |
Db     3404 TGATCTGGACCTAATGAATTCCTCAGGT 3431


RESULT 6
HSM130A
LOCUS       HSM130A                 3703 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen.
ACCESSION   Z22968
VERSION     Z22968.1  GI:312141
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3703)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
REFERENCE   2  (bases 1 to 3703)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
FEATURES             Location/Qualifiers
     source          1..3703
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3452
                     /codon_start=1
                     /product="M130 antigen"
                     /protein_id="CAA80541.1"
                     /db_xref="GI:312142"
                     /db_xref="SPTREMBL:Q07898"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSRYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVA
                     LSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQS
                     QTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICD
                     DSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGW
                     GQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVG
                     VVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSE
                     ETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALK
                     AFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDISVQK
                     TPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQ
                     IQYREMNSCLNADDLDLMNSSGGHSEPH"
     sig_peptide     102..221
     mat_peptide     222..3449
                     /product="M130 antigen"
                     /evidence=experimental
     polyA_signal    3677..3682
     polyA_signal    3681..3686
     polyA_signal    3685..3690
     polyA_site      3700
BASE COUNT      983 a    731 c   1059 g    930 t
ORIGIN

  Query Match            31.3%;  Score 1448.2;  DB 9;  Length 3703;
  Best Local Similarity  65.4%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches 1128;  Indels   24;  Gaps    3;

Qy     1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
              |||  ||  | |||  | || | |                | ||  |  |  |  .  ||
Db      107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy     1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
              ||| | | | | | | || ||||          | | |          |  |    . |  ||| |
Db      167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy     1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
              |||    ||| ||| | ||||  |||||      ||.|| |||    |||||||| || ||||
Db      227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy     1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
             | | || ||   |||||  | ||   | ||| |    |    |||||   |    |||   |   |||
Db      287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy     1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
             | ||||||| ||||| ||||||||          ||   |    |   |  |  ||||| |  ||
Db      347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy     1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
              |    |  || ||  ||||||||   |'      | |||||    ||||||||||||||||
Db      407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy     1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
             ||||| |||| | ||||||||     | |||||    | |   ||    |      |  .|||||
Db      467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy     1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
             |||||| |   || || |||      |   || ||||   | |||||      | |.    |
Db      527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy     1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538
             |||| ||| ||| | ||| | |||| ||||||| |||||| || |||||:||.| |
```

```
Db    587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db    647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy   1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db    707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1846

Qy   2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db   1847 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1906

Qy   2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db   1907 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966

Qy   2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db   1967 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 2026

Qy   2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db   2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 2086

Qy   2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db   2087 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 2146

Qy   3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
Db   2147 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 2206

Qy   3099 GGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
Db   2207 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 2266

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db   2267 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 2326

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   2327 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 2386

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   2387 TGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 2446

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGA 3398
Db   2447 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAA 2506

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   2507 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 2566

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   2567 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTACAATGGAGCTTG 2626

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   2627 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 2686

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   2687 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 2746

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   2747 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 2806
```

```
Qy   3699 TGCCCCATGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   2807 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 2866

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db   2867 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 2926

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
Db   2927 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 2986

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
Db   2987 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 3046

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
Db   3047 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 3106

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
Db   3107 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 3166

Qy   4059 GTGCTCTGGACAGTCGCTGAA-----------------ATCACTGAATGCCTCCTCAGG 4100
Db   3167 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 3226

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
Db   3227 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 3286

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
Db   3287 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAG 3343

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
Db   3344 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 3403

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGAC 4311
Db   3404 TGATCTGGACCTAATGAATTCCTCAGGAGGC 3434
```

```
RESULT 7
AX281768
LOCUS       AX281768                3707 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 177 from Patent WO0177389.
ACCESSION   AX281768
VERSION     AX281768.1  GI:16609019
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
            Mikita,T. and Tai,J.
  TITLE     Genes expressed in foam cell differentiation
  JOURNAL   Patent: WO 0177389-A 177 18-OCT-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3707
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 236574.12"
BASE COUNT      984 a    731 c   1062 g    930 t
ORIGIN

  Query Match             30.5%;  Score 1409.4;  DB 6;  Length 3707;
  Best Local Similarity   65.2%;  Pred. No. 0;
  Matches 2176;  Conservative    0;  Mismatches 1131;  Indels   28;  Gaps    6;
```

```
Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGT-TCAGGGAGAGTAGAGGTGA 1180
Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGGAGAGTGGAAGTGA 286

Qy   1181 GAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTG 1240
Db    287 AAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTG 346

Qy   1241 TGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTA 1300
Db    347 TGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGCTAATTCCA 406

Qy   1301 GTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTC 1360
Db    407 GTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTC 466

Qy   1361 TCTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATG 1417
Db    467 TTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATG 526

Qy   1418 CTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCC 1477
Db    527 CTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATA 586

Qy   1478 CCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACA 1537
Db    587 TGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATA 646

Qy   1538 GATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGC 1597
Db    647 ACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCA 706

Qy   1598 ATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTT 1657
Db    707 GTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTA 766

Qy   1658 CTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATA 1717
Db    767 TATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATA 826

Qy   1718 ATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGA 1777
Db    827 ACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGA 886

Qy   1778 GGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGT 1837
Db    887 GACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAAT 946

Qy   1838 GGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGC 1897
Db    947 GGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAAC 1006

Qy   1898 TGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAA 1957
Db   1007 TGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGAC 1066

Qy   1958 AAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGA 2017
Db   1067 ACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAAC 1126

Qy   2018 ACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTG 2077
Db   1127 ACCATGAATGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTG 1186
```

```
Qy   2078 ATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTG 2137
Db   1187 ATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTG 1246

Qy   2138 AGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTG 2197
Db   1247 AGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAG 1306

Qy   2198 CTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTC 2257
Db   1307 CTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGT 1366

Qy   2258 ATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAG 2317
Db   1367 ACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAA 1426

Qy   2318 CCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGG 2377
Db   1427 CTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAG 1486

Qy   2378 AAGCAAGTTTGATCTGCTCAG--CCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCC 2435
Db   1487 AAGCCAAAATTACCTGCTCAGGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCC 1546

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   1547 CTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCAGA 1606

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
Db   1607 CTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTC 1666

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCA 2615
Db   1667 TATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCA 1726

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   1727 GTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAAC 1786

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735
Db   1787 TTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGT 1846

Qy   2736 GAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTC 2795
Db   1847 GAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATC 1906

Qy   2796 ACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTG 2855
Db   1907 CCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATG 1966

Qy   2856 TGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTG 2915
Db   1967 TGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTG 2026

Qy   2916 GGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGT 2975
Db   2027 GAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGC 2086

Qy   2976 TCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCT 3035
Db   2087 TCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCA 2146

Qy   3036 GACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCC 3095
Db   2147 GTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCC 2206

Qy   3096 AGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAG 3155
Db   2207 AGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGG 2266

Qy   3156 CCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGA 3215
Db   2267 TCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGA 2326

Qy   3216 CGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTT 3275
Db   2327 CAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCAT 2386

Qy   3276 CAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCT 3335
Db   2387 TAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGAT 2446

Qy   3336 GAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCA 3395
Db   2447 GAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCA 2506

Qy   3396 CGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCT 3455
Db   2507 AAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACT 2566

Qy   3456 CTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGAC 3515
Db   2567 GACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGC 2626

Qy   3516 CTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCA 3575
Db   2627 TTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCA 2686

Qy   3576 GCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGG 3635
Db   2687 GCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCAT 2746

Qy   3636 TTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATAT-GGCAGTGCC 3694
Db   2747 TCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGGCAGTGCC 2806

Qy   3695 TGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTG 3754
Db   2807 CATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTG 2866

Qy   3755 AAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGC 3814
Db   2867 ACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGC 2926

Qy   3815 ACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGG 3874
Db   2927 ATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGG 2986

Qy   3875 TGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCC 3934
Db   2987 TGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTC 3046

Qy   3935 AGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTAT 3994
Db   3047 AGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGT 3106

Qy   3995 GGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCG 4054
Db   3107 GGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAG 3166

Qy   4055 TGAGGTGCTCTGGACAGTCGCTGAA-----------------ATCACTGAATGCCTCCT 4096
Db   3167 TGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCAT 3226

Qy   4097 CAGGTCATTTAGCACTTATTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTA 4156
Db   3227 CCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCG 3286

Qy   4157 TTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAA 4216
Db   3287 TCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCT 3343

Qy   4217 CCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGA 4276
```

```
Db    3344 CAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATG 3403

Qy    4277 GAGAGGACCCACATGGACAAGAACCTCAGATGAC 4311

Db    3404 CAGATGATCTGGACCTAATGAATTCCTCAGGAGGC 3438


RESULT 8
HSM130AE
LOCUS       HSM130AE                3802 bp    mRNA    linear   PRI 22-NOV-1996
DEFINITION  H.sapiens mRNA for M130 antigen extracellular variant.
ACCESSION   Z22971
VERSION     Z22971.1  GI:312147
KEYWORDS    antigen; antigen M130.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3802)
  AUTHORS   Micklem,K.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
            of Pathology, University of Oxford, Level 1, Maternity Block, John
            Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE   2  (bases 1 to 3802)
  AUTHORS   Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
            and Mason,D.Y.
  TITLE     A new macrophage differentiation antigen which is a member of the
            scavenger receptor superfamily
  JOURNAL   Eur. J. Immunol. 23 (9), 2320-2325 (1993)
  MEDLINE   93380506
   PUBMED   8370408
FEATURES             Location/Qualifiers
     source          1..3802
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Macrophage"
     CDS             102..3551
                     /codon_start=1
                     /product="M130 antigen extracellular variant"
                     /protein_id="CAA80544.1"
                     /db_xref="GI:312148"
                     /db_xref="SPTREMBL:Q07901"
                     /translation="MVLLEDSGSADFRRHFVNLSPFTITVVLLLSACFVTSSLGGTDK
                     ELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWANSS
                     AGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG
                     NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWF
                     DDLICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEV
                     RFQGEWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHE
                     PAVWQCKHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKV
                     CDRGWGLKEADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNW
                     QWGGLTCDHYEEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAAS
                     VLCRELQCGTVVSILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSR
                     DVGVVCSSKTQKTSLIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRV
                     ELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTG
                     TEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVA
                     CIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATG
                     SAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTS
                     EASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMSI
                     PMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEI
                     WHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNE
                     SSLWDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVV
                     LLAIFVALFFLTKKRRQRQRLAVSSRGENLVHQIQYREMNSCLNADDLDLMNSSGGHS
                     EPH"
     sig_peptide     102..221
     mat_peptide     222..3548
                     /product="M130 antigen extracellular variant"
                     /evidence=experimental
     polyA_signal    3776..3781
     polyA_signal    3780..3785
     polyA_signal    3784..3789
     polyA_site      3799
BASE COUNT     1009 a    751 c    1081 g    961 t
ORIGIN

  Query Match             28.9%;  Score 1339.2;  DB 9;  Length 3802;
  Best Local Similarity   63.5%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches 1128;  Indels  123;  Gaps    4;

Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061

Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121

Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181

Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy   1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241

Db    287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy   1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301

Db    347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy   1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361

Db    407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy   1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418

Db    467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy   1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478

Db    527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy   1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538

Db    587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598

Db    647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy   1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658

Db    707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718

Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778

Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838

Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898

Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958

Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018

Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126
```

```
Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTG----------------------------- 2708
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGTAAGACCCAGAAAACATCTTTAAT 1846

Qy   2709 ------------------------------------------------------------ 2708
Db   1847 TGGTTCTTATACTGTGAAAGGGACAGGGTTAGGGAGTCATAGCTGTCTTTTTCTAAAGCC 1906

Qy   2709 ---------TTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGG 2759
Db   1907 CTGTCTCCTTCCAGGATACACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGTGAGGG 1966

Qy   2760 GCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCC 2819
Db   1967 CAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCTCTGTAACTCTCACTGGGACAT 2026

Qy   2820 AGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGG 2879
Db   2027 AGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACCCCAGG 2086

Qy   2880 AGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGG 2939
Db   2087 AGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAGGCATATGTTTCACTGCACTGG 2146

Qy   2940 GAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCA 2999
Db   2147 GACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCTAGGTGCTTCATTATGTCCTTC 2206

Qy   3000 TGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCT 3059
Db   2207 AGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTCCCAAACACTGTCCTCGTGCAA 2266

Qy   3060 CGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTT 3119
Db   2267 TTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCCTGCAT 2326

Qy   3120 AGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGAT 3179
Db   2327 AGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTAGAGAT 2386

Qy   3180 CTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCA 3239
Db   2387 CTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAGTGATGCCCA 2446

Qy   3240 CGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTT 3299
Db   2447 CGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGGTTCTGCTCATTT 2506

Qy   3300 TGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCA 3359
Db   2507 TGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAAATGCAATGGAAAAGAATCCCG 2566

Qy   3360 CTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGC 3419
Db   2567 CATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAATTGCAGGCACAAGGAGGATGC 2626

Qy   3420 AGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAG 3479
Db   2627 GGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGACCAGTGAAGCCAGCAGAGAGGC 2686

Qy   3480 CTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAA 3539
Db   2687 CTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTGGGGCACTGTTGGCAAGAGTAG 2746

Qy   3540 CATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGT 3599
Db   2747 CATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCAGACAAAGGGAA 2806

Qy   3600 TGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCA 3659
Db   2807 AATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTGGACAATGTTCA 2866

Qy   3660 GTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAAT 3719
Db   2867 GTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGGGAGAAGAGACT 2926

Qy   3720 CTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGG 3779
Db   2927 GGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAACAAGATAAGACTTCAGGAAGG 2986

Qy   3780 AGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTG 3839
Db   2987 ACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGAGGTTCCTGGGGGACAGTGTG 3046

Qy   3840 TGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTC 3899
Db   3047 TGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTTGGCTGTGGTCC 3106

Qy   3900 TGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGA 3959
Db   3107 AGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCGATATGGCTCAA 3166

Qy   3960 TGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGG 4019
Db   3167 TGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGATTGTCCTGCCAGACGCTGGGG 3226

Qy   4020 ACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAA 4079
Db   3227 CCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAATTGCACAGATATTTCAGTGCA 3286

Qy   4080 ------------------ATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATC 4121
```

```
Db    3287 GAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCGTCAGTCATCCTTTATTGCAGT 3346

Qy    4122 CAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGT 4181

Db    3347 CGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGCATTATTCTTCTTGACTAAAAA 3406

Qy    4182 TCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGA 4241

Db    3407 GCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAGAGGAGAGAACTTAGTCCACCA 3463

Qy    4242 GAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAAC 4301

Db    3464 AATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGATGATCTGGACCTAATGAATTC 3523

Qy    4302 CTCAGATGAC 4311

Db    3524 CTCAGGAGGC 3533


RESULT 9
AF274883
LOCUS       AF274883                4405 bp    mRNA    linear   ROD 10-MAY-2001
DEFINITION  Mus musculus macrophage hemoglobin scavenger receptor CD163
            precursor, mRNA, complete cds.
ACCESSION   AF274883
VERSION     AF274883.1  GI:13236130
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4405)
  AUTHORS   Schaer,D.J., Boretti,F.S., Hongegger,A., Poehler,D., Linnscheid,P.,
            Staege,H., Muller,C., Schoedon,G. and Schaffner,A.
  TITLE     Molecular cloning and characterization of the mouse CD163
            homologue, a highly glucocorticoid-inducible member of the
            scavenger receptor cysteine-rich family
  JOURNAL   Immunogenetics 53 (2), 170-177 (2001)
  MEDLINE   21243729
   PUBMED   11345593
REFERENCE   2  (bases 1 to 4405)
  AUTHORS   Schaer,D.J., Boretti,F.S. and Schaffner,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2000) Departement of Medicine, Medizinische
            Klinik B, University Hospital Zurich, Ramistrasse 100, Zurich, ZH
            8091, Switzerland
FEATURES             Location/Qualifiers
     source          1..4405
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="6"
                     /map="15.43 centirays distal from the D6Mit300"
     CDS             85..3450
                     /note="FAT; member of the scavenger receptor cysteine rich
                     type B superfamily; transmembrane protein type I"
                     /codon_start=1
                     /product="macrophage hemoglobin scavenger receptor CD163
                     precursor"
                     /protein_id="AAK16065.1"
                     /db_xref="GI:13236131"
                     /translation="MGGHRMVLLGGAGSPGCKRFVHLGFFVVAVSSLLSASAVTNAPG
                     EMKKELRLAGGENNCSGRVELKIHDKWGTVCSNGWSMNEVSVVCQQLGCPTSIKALGW
                     ANSSAGSGYIWMDKVSCTGNESALWDCKHDGWGKHNCTHEKDAGVTCSDGSNLEMRLV
                     NSAGHRCLGRVEIKFQGKWGTVCDDNFSKDHASVICKQLGCGSAISFSGSAKLGAGSG
                     PIWLDDLACNGNESALWDCKHRGWGKHNCDHAEDVGVICLEGADLSLRLVDGVSRCSG
                     RLEVRFQGEWGTVCDDNWDLRDASVVCKQLGCPTAISAIGRVNASEGSGQIWLDNISC
                     EGHEATLWECKHQEWGKHYCHHREDAGVTCSDGADLELRLVGGGSRCAGIVEVEIQKL
                     TGKMCSRGWTLADADVVCRQLGCGSALQTQAKIYSKTGATNTWLFPGSCNGNETTFWQ
                     CKNWQWGGLSCDNFEEAKVTCSGHREPRLVGGEIPCSGRVEVKHGDVWGSVCDFDLSL
                     EAASVVCRELQCGTVVSILGGAHFGEGSGQIWGEEFQCSGDESHLSLCSVAPPLDRTC
                     THSRDVSVVCSRYIDIRLAGGESSCEGRVELKTLGAWGPLCSSHWDMEDAHVLCQQLK
                     CGVAQSIPEGAHFGKGAGQVWSHMFHCTGTEEHIGDCLMTALGAPTCSEGQVASVICS
                     GNQSQTLLPCSSLSPVQTTSSTIPKESEVPCIASGQLRLVGGGGRCAGRVEVCHEGSW
                     GTVCDDNWDMTDANVVCKQLDCGVAINATGSAYFGEGAGAIWLDEVICTGKESHIWQC
                     HSHGWGRHNCRHKEDAGVICSEFMSLRLTNEAHKENCTGRLEVFYNGTWGSIGSSNMS
                     PTTVGVVCRQLGCADNGTVKPIPSDKTPSRPMWVDRVQCPKGVDTLWQCPSSPWKQRQ
                     ASPSSQESWIICDNKIRLQEGHTDCSGRVEIWHKGSWGTVCDDSWDLNDAKVVCKQLG
                     CGQAVKALKEAAFGPGTGPIWLNEIKCRGNESSLWDCPAKPWSHSDCGHKEDASIQCL
                     PKMTSESHHGTGHPTLTALLVCGAILLVLLIVFLLWTLKRRRIQRLTVSSRGEVLIHQ
                     VQYQEMDSKADDLDLLKSSGVIQRHTEKENDNL"
     sig_peptide     85..198
     misc_feature    3220..3288
                     /note="Region: transmembrane domain"
BASE COUNT     1174 a    884 c   1246 g   1101 t
ORIGIN

  Query Match           27.5%;  Score 1270.4;  DB 10;  Length 4405;
  Best Local Similarity 63.6%;  Pred. No. 3.8e-306;
  Matches 2001;  Conservative    0;  Mismatches 1131;  Indels   12;  Gaps    4;

Qy    1128 GGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCA 1187

Db     228 GGAACTGAGACTGGCGGGTGGTGAAAACAACTGTAGTGGGAGAGTGGAACTTAAGATCCA 287

Qy    1188 TGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTG 1247

Db     288 TGACAAGTGGGGCACAGTGTGCAGTAACGGCTGGAGCATGAATGAAGTGTCCGTGGTTTG 347

Qy    1248 TAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGA 1307

Db     348 CCAGCAGCTGGGATGCCCAACTTCTATTAAAGCCCTTGGATGGGCTAACTCCAGCGCCGG 407

Qy    1308 AGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGA 1367

Db     408 CTCTGGATATATCTGGATGGACAAAGTTTCTTGTACAGGGAATGAGTCAGCTCTTTGGGA 467

Qy    1368 CTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAAT 1427

Db     468 CTGCAAACATGATGGGTGGGGAAAGCATAACTGTACCCATGAAAAAGATGCTGGAGTGAC 527

Qy    1428 TTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGT---CGGGGCTCATAGCCCCTGTTA 1484

Db     528 CTGCTCAGATGGATCTAATTTGGAGATGAGACTGGTGAACAGTGCGGGCCACCGATGCTT 587

Qy    1485 TGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGACAGATGGAG 1544

Db     588 AGGAAGAGTAGAAATAAAGTTCCAGGGAAAGTGGGGACGGTGTGTGACGACAACTTCAG 647

Qy    1545 CACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTT 1604

Db     648 CAAAGATCACGCTTCTGTGATTTGTAAACAGCTTGGATGTGGAAGTGCCATTAGTTTCTC 707

Qy    1605 TGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCAT 1664

Db     708 TGGCTCAGCTAAATTGGGAGCTGGTTCTGGACCAATCTGGCTCGATGACCTGGCATGCAA 767

Qy    1665 TGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGT 1724

Db     768 TGGAAATGAGTCAGCTCTCTGGGACTGCAAACACCGGGGATGGGGCAAGCATAACTGTGA 827

Qy    1725 ACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGT 1784

Db     828 CCATGCTGAGGATGTCGGTGTGATTTGCTTAGAGGGAGCAGATCTGAGCCTGAGACTAGT 887

Qy    1785 GGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCAC 1844

Db     888 GGATGGAGTGTCCAGATGTTCAGGAAGATTGGAAGTGAGATTCCAAGGAGAATGGGGGAC 947

Qy    1845 AGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTG 1904

Db     948 CGTGTGTGATGATAACTGGGATCTCCGGGATGCTTCTGTGGTGTGCAAGCAACTGGGATG 1007

Qy    1905 CCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTG 1964

Db    1008 TCCAACTGCCATCAGTGCCATTGGTCGAGTTAATGCCAGTGAGGGATCTGGACAGATTTG 1067
```

```
Qy   1965 GCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGG 2024
Db   1068 GCTTGACAACATTTCATGCGAAGGACATGAGGCAACTCTTTGGGAGTGTAAACACCAAGA 1127

Qy   2025 GTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATC 2084
Db   1128 GTGGGGAAAGCATTACTGTCATCATAGAGAAGACGCTGGTGTGACATGTTCTGATGGAGC 1187

Qy   2085 GGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAA 2144
Db   1188 AGATCTGGAACTTAGACTTGTAGGTGGAGGCAGTCGCTGTGCTGGCATTGTGGAGGTGGA 1247

Qy   2145 TGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGT 2204
Db   1248 GATTCAGAAGCTGACTGGGAAGATGTGTAGCCGAGGCTGGACACTGGCAGATGCGGATGT 1307

Qy   2205 TGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCAC 2264
Db   1308 GGTTTGCAGACAGCTTGGATGTGGATCTGCGCTTCAAACCCAGGCTAAGATCTACTCTAA 1367

Qy   2265 AGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCT 2324
Db   1368 AACTGGGGCAACAAATACGTGGCTCTTTCCTGGATCTTGTAATGGAAATGAAACTACTTT 1427

Qy   2325 CTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAG 2384
Db   1428 TTGGCAATGCAAAAACTGGCAGTGGGGCGGCCTTTCCTGTGATAATTTCGAAGAAGCCAA 1487

Qy   2385 TTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGG 2444
Db   1488 AGTTACCTGCTCAGGCCACAGGGAACCCAGACTGGTTGGAGGAGAAATCCCATGCTCTGG 1547

Qy   2445 ACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCT 2504
Db   1548 TCGTGTGGAAGTGAAACACGGAGACGTGTGGGGCTCCGTCTGTGATTTTGACTTGTCTCT 1607

Qy   2505 TCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGT 2564
Db   1608 GGAAGCTGCCAGTGTGGTGTGCAGGGAATTACAATGTGGAACAGTCGTCTCTATCCTAGG 1667

Qy   2565 GGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGG 2624
Db   1668 GGGAGCACATTTTGGAGAAGGAAGTGGACAGATCTGGGGTGAAGAATTCCAGTGTAGTGG 1727

Qy   2625 GAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCA 2684
Db   1728 GGATGAGTCCCATCTTTCACTATGCTCAGTGGCGCCCCCGCTAGACAGAACTTGTACCCA 1787

Qy   2685 CAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAA 2744
Db   1788 CAGCAGGGATGTCAGCGTAGTCTGCTCACGATACATAGATATTCGTCTGGCAGGCGGCGA 1847

Qy   2745 ATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGA 2804
Db   1848 GTCCTCCTGTGAGGGAAGAGTGGAGCTCAAGACACTCGGAGCCTGGGGTCCCCTCTGCAG 1907

Qy   2805 CACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGC 2864
Db   1908 TTCTCATTGGGACATGGAAGATGCTCATGTCTTATGTCAGCAGCTGAAGTGTGGGGTTGC 1967

Qy   2865 TCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAG 2924
Db   1968 CCAATCTATTCCAGAAGGAGCACATTTTGGGAAAGGAGCTGGTCAGGTCTGGAGTCACAT 2027

Qy   2925 GTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGC 2984
Db   2028 GTTCCACTGCACTGGAACTGAGGAACATATAGGAGATTGCCTCATGACTGCTCTGGGTGC 2087

Qy   2985 ACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCC 3044
Db   2088 GCCGACGTGTTCCGAAGGACAGGTGGCCTCTGTCATCTGCTCAGGAAACCAATCCCAGAC 2147

Qy   3045 ACTGTTTCCATGCCTCGCAAATGTATCTGACC---CATATTTGTCTGCAGTTCCAGAGGG 3101
Db   2148 ACTATTGCCATGTAGTTCATTGTCTCCAGTCCAAACAACAAGCTCTACAATTCCAAAGGA 2207

Qy   3102 CAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTG 3161
Db   2208 GAGTGAAGTTCCCTGCATAGCAAGTGGCCAGCTTCGCTTGGTAGGTGGAGGTGGTCGCTG 2267

Qy   3162 TGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTG 3221
Db   2268 CGCTGGAAGAGTGGAGGTCTGCCACGAGGGCTCTTGGGGCACCGTCTGTGATGACAATTG 2327

Qy   3222 GGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGC 3281
Db   2328 GGATATGACTGATGCCAATGTGGTGTGCAAGCAGCTGGACTGTGGCGTGGCAATTAACGC 2387

Qy   3282 CACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGCCCATCTGGCTGGATGACCTGAACTG 3341
Db   2388 CACTGGCTCTGCTTACTTCGGGGAAGGAGCAGGAGCTATCTGGCTAGACGAAGTCATCTG 2447

Qy   3342 CACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTG 3401
Db   2448 CACTGGGAAAGAGTCTCATATTTGGCAGTGCCATTCACATGGCTGGGGACGCCATAACTG 2507

Qy   3402 CAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAG 3461
Db   2508 CAGGCACAAAGAAGATGCAGGTGTTATCTGCTCCGAGTTCATGTCTCTGAGGCTGACCAA 2567

Qy   3462 TGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGG 3521
Db   2568 CGAAGCCCACAAAGAAAACTGCACAGGTCGCCTTGAAGTGTTTTACAATGGTACATGGGG 2627

Qy   3522 CAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGG 3581
Db   2628 CAGTATTGGCAGTAGCAATATGTCTCCAACCACTGTGGGGGTGGTGTGCCGTCAGCTGGG 2687

Qy   3582 CTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCAT 3641
Db   2688 CTGTGCAGACAACGGGACTGTGAAACCCATACCTTCAGACAAGACACCATCCAGGCCCAT 2747

Qy   3642 GTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGC 3701
Db   2748 GTGGGTAGATCGTGTGCAGTGTCCAAAAGGAGTTGACACTTTGTGGCAGTGCCCCTCGTC 2807

Qy   3702 CCCATGGGAGCGAAGAATCTCCAGCC---CAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   2808 ACCTTGGAAACAGAGACAGGCCAGCCCCTCCTCCCAGGAGTCCTGGATCATCTGTGACAA 2867

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db   2868 CAAAATAAGACTCCAGGAAGGGCATACAGACTGTTCTGGACGTGTGGAGATCTGGCACAA 2927

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
Db   2928 AGGTTCCTGGGGAACAGTGTGTGATGACTCCTGGGATCTTAATGATGCTAAGGTTGTATG 2987

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
Db   2988 TAAGCAGTTGGGCTGTGGCCAAGCTGTGAAGGCACTAAAAGAAGCAGCATTTGGTCCAGG 3047

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
Db   3048 AACTGGGCCCATATGGCTCAATGAAATTAAGTGTAGAGGGAATGAGTCTTCCCTGTGGGA 3107

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
Db   3108 TTGTCCTGCCAAACCGTGGAGTCACAGCGACTGTGGGCACAAAGAAGATGCTTCCATCCA 3167

Qy   4059 GTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTT 4118
Db   3168 GTGCCTCCCAAAAATGACTTCAGAATCACATCATGGCACAGGTCACCCCACCCTCACGGC 3227

Qy   4119 ATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCG 4178
```

```
Db     3228 ACTCTTGGTTTGTGGAGCCATTCTATTGGTCCTCCTCATTGTCTTCCTCCTGTGGACTC- 3286

Qy     4179 AGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGA 4238

Db     3287 --TGAAGCGACGACGGATTCAGCGACTTACAGTTTCCTCAAGAGGAGAGGTCTTGATACA 3344

Qy     4239 GGAGAATTTATTCCATGAGATGGA 4262

Db     3345 TCAAGTTCAGTACCAAGAGATGGA 3368


RESULT 10
BBWC11MR
LOCUS       BBWC11MR                4475 bp    mRNA    linear   MAM 24-MAY-1994
DEFINITION  B.bovis WC1.1 mRNA.
ACCESSION   X63723
VERSION     X63723.1  GI:13
KEYWORDS    WC1.1 protein.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 4475)
  AUTHORS   Wijngaard,P.L., MacHugh,N.D., Metzelaar,M.J., Romberg,S.,
            Bensaid,A., Pepin,L., Davis,W.C. and Clevers,H.C.
  TITLE     Members of the novel WC1 gene family are differentially expressed
            on subsets of bovine CD4-CD8- gamma delta T lymphocytes
  JOURNAL   J. Immunol. 152 (7), 3476-3482 (1994)
  MEDLINE   94194107
   PUBMED   7511649
REFERENCE   2  (bases 1 to 4475)
  AUTHORS   Wijngaard,P.L.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-1992) P.L.J. Wijngaard, Dept of Immunology,
            University Hospital, P O Box 85500, 3508 GA Utrecht, THE
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1..4475
                     /organism="Bos taurus"
                     /strain="indicus"
                     /isolate="D409"
                     /db_xref="taxon:9913"
                     /clone="pBoWC1.1"
                     /cell_line="D409/N2"
                     /tissue_type="blood"
                     /clone_lib="D409/N2.H1"
                     /dev_stage="gamma/delta T-cell"
     gene            22..4332
                     /gene="WC1.1 subcluster gene"
     CDS             22..4332
                     /gene="WC1.1 subcluster gene"
                     /codon_start=1
                     /product="BoWC1.1"
                     /protein_id="CAA45255.1"
                     /db_xref="GI:14"
                     /db_xref="SWISS-PROT:P30205"
                     /translation="MALGRHLSLRGLCVLLLGTMVGGQALELRLKDGVHRCEGRVEVK
                     HQGEWGTVDGYRWTLKDASVVCRQLGCGAAIGFPGGAYFGPGLGPIWLLYTSCEGTES
                     TVSDCEHSNIKDYRNDGYNHGRDAGVVCSGFVRLAGGDGPCSGRVEVHSGEAWIPVSD
                     GNFTLATAQIICAELGCGKAVSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPC
                     PGGTCHHSGSAQVVCSAYSEVRLMTNGSSQCEGQVEMNISGQWRALCASHWSLANANV
                     ICRQLGCGVAISTPGGPHLVEEGDQILTARFHCSGAESFLWSCPVTALGGPDCSHGNT
                     ASVICSGNQIQVLPQCNDSVSQPTGSAASEDSAPYCSDSRQLRLVDGGGPCAGRVEIL
                     DQGSWGTICDDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKES
                     HVWRCPSRGWGQHNCRHKQDAGVICSEFLALRMVSEDQQCAGWLEVFYNGTWGSVCRN
                     PMEDITVSTICRQLGCGDSGTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWN
                     YNSCSPKEEAYIWCADSRQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVC
                     KQLGCGEALDATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAG
                     VICSGFVRLAGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLG
                     HMPFRESDGQVWAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMK
                     NGTSQCEGQVEMKISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQ
```

```
                     ISTAQFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAG
                     SAASEESSPYCSDSRQLRLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQL
                     GCGEALNATGSAHFGAGSGPIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVIC
                     SEFLALRMVSEDQQCAGWLEVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTS
                     VGLREGSRPRWVDLIQCRKMDTSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTA
                     AACTDREKLRLRGGDSECSGRVEVWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEA
                     VRSAAFGPGNGSIWLDEVQCGGRESSLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLP
                     TTTAGTRTTSNSLPGIFSLPGVLCLILGSLLFLVLVILVTQLLRWRAERRALSSYEDA
                     LAEAVYEELDYLLTQKEGLGSPDQMTDVPDENYDDAEEVPVPGTPSPSQGNEEEVPPE
                     KEDGVRSSQTGSFLNFSREAANPGEGEESFWLLQGKKGDAGYDDVELSALGTSPVTFS
                     "
     polyA_site      4453
BASE COUNT      933 a   1113 c   1442 g    987 t
ORIGIN

  Query Match             19.2%;  Score 886.8;  DB 4;  Length 4475;
  Best Local Similarity   54.9%;  Pred. No. 2.8e-210;
  Matches 2061;  Conservative    0;  Mismatches 1527;  Indels  168;  Gaps    9;

Qy      483 TCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATT 542

Db       96 TCTGGAGCTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCA 155

Qy      543 CCAAGAAAGGTGGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGT 602

Db      156 CCAAGGAGAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGT 215

Qy      603 GTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGC 662

Db      216 GTGCAGACAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTTGGGCC 275

Qy      663 TGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTG 722

Db      276 AGGACTTGGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTCAG 335

Qy      723 GAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATT 782

Db      336 TGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGGGA 395

Qy      783 AACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGTAT 842

Db      396 TGCTGGAGTAGTCTGCTCAGGATTTG---TGCGTCTGGCTGGAGGGGATGGACCCTGCTC 452

Qy      843 GGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGGACCGTATGCCACCATAAGTGGAA 902

Db      453 AGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGATGGGAACTTCAC 512

Qy      903 CAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTCGC 962

Db      513 ACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCTGTGTCTGTCCT 572

Qy      963 TGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTC 1022

Db      573 GGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAGTTCAGGTGTGA 632

Qy     1023 CGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGTCT 1082

Db      633 GGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGGGGCACGTGTCA 692

Qy     1083 TCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTAGC 1142

Db      693 CCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGA---AGTCCGGCTCATGAC 749

Qy     1143 AGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGAC 1202

Db      750 AAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGACAATGGAGAGC 809

Qy     1203 AATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATG 1262

Db      810 GCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGTCAGCTCGGCTG 869

Qy     1263 TCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTG 1322
```

```
Db    870 TGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGTGATCAGATCCT 929

Qy   1323 GATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGG 1382
Db    930 AACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGACTGC 989

Qy   1383 AAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGC 1442
Db    990 CCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAACCA 1049

Qy   1443 AGATCTGG------------------------------------------------------ 1450
Db   1050 GATCCAGGTGCTTCCCCAGTGCAACGACTCCGTGTCTCAACCTACAGGCTCTGCGGCCTC 1109

Qy   1451 -----------------------------------ACCTAAGGCTTGTCGGGGCTCATAG 1475
Db   1110 AGAGGACAGCGCCCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGGCGG 1169

Qy   1476 CCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGA 1535
Db   1170 TCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGA 1229

Qy   1536 CAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTAT 1595
Db   1230 CGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGCCCT 1289

Qy   1596 GCATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGT 1655
Db   1290 CAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGACAACTT 1349

Qy   1656 TTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCA 1715
Db   1350 GAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCAGCA 1409

Qy   1716 TAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCT 1775
Db   1410 CAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCAGAGT------TCCTGGCCCT 1463

Qy   1776 GAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACG 1835
Db   1464 CAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAATGGGAC 1523

Qy   1836 GTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCA 1895
Db   1524 CTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGCAGACA 1583

Qy   1896 GCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGG 1955
Db   1584 GCTTGGCTGTGGGGACAGTGGAACCCTCAACTCTTCTGTTGCTCTTAGAGAAGGTTTTAG 1643

Qy   1956 AAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAG 2015
Db   1644 GCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACCTCTCTCTGGCAGTGTCC 1703

Qy   2016 GAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTC 2075
Db   1704 TTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAAGCCTATATCTGGTG--- 1760

Qy   2076 TGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGT 2135
Db   1761 ---TGCAGACAGCAGACAGATCCGCCTGGTGGATGGAGGTGGTCGCTGCTCTGGGAGAGT 1817

Qy   2136 TGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACAT 2195
Db   1818 GGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGACCGCTGGGACCTGGACGA 1877

Qy   2196 TGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCC 2255
Db   1878 TGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTGGACGCCACTGTCTCTTC 1937

Qy   2256 TCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGA 2315
Db   1938 CTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTGAACTGCAGAGGAGAGGA 1997

Qy   2316 AGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATAT 2375
Db   1998 GTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACACAACTGCAATCATCAAGA 2057

Qy   2376 GGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCC 2435
Db   2058 AGATGCAGGAGTCATCTGCTCAGGATTTGTGC------GTCTGGCTGGAGGAGATGGACC 2111

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   2112 CTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACCCCAGTGTCTGATGGAAA 2171

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
Db   2172 CTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGATGTGGCAAGGCTGTGTC 2231

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCA 2615
Db   2232 TGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTCTGGGCTGAAGAGTTCAG 2291

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   2292 GTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTGCCCTGTCCAGGAGGCAC 2351

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735
Db   2352 ATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTACACAGAAGTCCAGCTTAT 2411

Qy   2736 G---AATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGG 2792
Db   2412 GAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAGATCTCTGGACGATGGAG 2471

Qy   2793 CTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAG 2852
Db   2472 AGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTGTCTGTCGTCAGCTCGG 2531

Qy   2853 CTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGT 2912
Db   2532 CTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTGGAAGGAGGTGATCAGAT 2591

Qy   2913 GTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGAC 2972
Db   2592 CTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGAC 2651

Qy   2973 AGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAG 3032
Db   2652 TGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAA 2711

Qy   3033 CCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGT 3092
Db   2712 CCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAACCTGCAGGCTCTGCGGC 2771

Qy   3093 TCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGA 3152
Db   2772 CTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGG 2831

Qy   3153 CAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGA 3212
Db   2832 CGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGA 2891

Qy   3213 TGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGC 3272
Db   2892 TGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGC 2951

Qy   3273 CTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGA 3332
Db   2952 CCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGCTGGACGA 3011

Qy   3333 CCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCA 3392
Db   3012 CCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCG 3071
```

```
Qy   3393 GCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAG 3452
Db   3072 GCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCAGAGTTCCTGGCCCTCAG 3131

Qy   3453 GCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGG 3512
Db   3132 G------ATGGTGAGCGAGGACCAGCAGTGTGCTGGGTGGCTGGAGGTTTTCTACAACGG 3185

Qy   3513 GACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAG 3572
Db   3186 GACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACTGTGTCCGTGATCTGCAG 3245

Qy   3573 GCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTC 3632
Db   3246 ACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTTGGTCTCAGGGAAGGTTC 3305

Qy   3633 TGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTG 3692
Db   3306 TAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGATACCTCTCTCTGGCAGTG 3365

Qy   3693 CCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATG 3752
Db   3366 TCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAGGAAGCCTACATCTCATG 3425

Qy   3753 TGAAGATAGAATAA---------------------------------------------- 3766
Db   3426 TGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGCACAGACAGAGAGAAGCT 3485

Qy   3767 --GAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTC 3824
Db   3486 CCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAGGTGTGGCACAACGGCTC 3545

Qy   3825 CTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCA 3884
Db   3546 CTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCTGAGGTGGTGTGTCAGCA 3605

Qy   3885 GCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGG 3944
Db   3606 GCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCATTTGGCCCTGGAAATGG 3665

Qy   3945 AACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCA 4004
Db   3666 GAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCCTCCCTGTGGGACTGTGT 3725

Qy   4005 CGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTC 4064
Db   3726 TGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGATGCTGGTGTGAGGTGCTC 3785

Qy   4065 TGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4100
Db   3786 TGGTGTAAGGACAACATTGCCCACGACCACAGCAGG 3821
```

```
RESULT 11
AF400115
LOCUS       AF400115                2305 bp    mRNA    linear   MAM 21-AUG-2001
DEFINITION  Ovis aries EM4b mRNA, partial cds.
ACCESSION   AF400115
VERSION     AF400115.1  GI:15217033
KEYWORDS    .
SOURCE      Ovis aries.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 2305)
  AUTHORS   Gill,J.W. and Walker,I.D.
  TITLE     Novel T19 transcripts from ovine CD4+ T cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2305)
  AUTHORS   Gill,J.W. and Walker,I.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2001) Veterinary Science, University of
            Melbourne, Cnr Park Drive and Flemington Rd, Parkville, VIC 3052,
            Australia
FEATURES             Location/Qualifiers
     source          1..2305
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
                     /cell_type="CD4+ T cell"
                     /tissue_type="efferent mesenteric lymph"
     CDS             <1..2133
                     /note="T19 multigene family member"
                     /codon_start=1
                     /product="EM4b"
                     /protein_id="AAK92469.1"
                     /db_xref="GI:15217034"
                     /translation="DWTPVSDGNFTFPTAQVICAELGCGNVVSVLGHVPFRESDGLVW
                     AEEFRCKGEEPELWSCPRVPCPGGTCHHSGAVHVVCSEVRLMKNGSSQCEGQVEMNIS
                     GRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHFVEGSDQISTVRFHCSGVESFL
                     WSCPVTALGGPDCSHGNMASVICSGNQTQVLPQCNDSMSERAGSVVSEEIAPYCSDSR
                     QLRLVDGDSPCAGRVEILDQGSWGTICDDGWDLDDARVVCRQLGCGEALNATGSAHFG
                     AGSGPIWLNDLNCTGNESHVWNCPSRGWGRHDCRHKEDAGVICSEFLALRMVSEDQQC
                     AGWLEVFYNGTWGSVCRSPMEDITVSIICSQLGCGDSGSLNSSVALREGFRPRWVDGI
                     RCRNTDTSLWQCPSDHWNYTSCSPKEEAYISCAGSRQLRLVDGGGPCAGRVEILDQGS
                     WGTICDDSWDLDDARVVCRQLGCGGALNALQFAGLGQGSGPIWLDELNCGGKESHVWR
                     CPSQGWGRHDCGHAEDAGIVCSGFVRLAGGEGPCSGRVEVHSGEDWTPVSDGNFTFPT
                     AQVICAELRCGKAVSVLGHVPFRESDGRVWAEEFRCEGEEPKLWSCPRVPCPGGTCHH
                     SGAAQVVCSVYAEVRLMTNGSSQCEGQVKEYGMRSSQKRSSLNVSREEANTGEGEESP
                     WLLQGDKGDPEYDDVELSVQGTSPVTFW"
BASE COUNT      512 a    544 c    716 g    533 t
ORIGIN

  Query Match             14.0%;  Score 649.8;  DB 4;  Length 2305;
  Best Local Similarity   64.7%;  Pred. No. 4.7e-151;
  Matches 1039;  Conservative    0;  Mismatches  542;  Indels   24;  Gaps    4;
```

```
Qy   2473 TGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAA 2532
Db      4 TGGACCCCAGTGTCTGATGGGAACTTTACATTCCCCACTGCACAGGTCATCTGTGCAGAA 63

Qy   2533 TTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGT 2592
Db     64 CTGGGATGTGGCAATGTTGTGTCTGTCCTGGGACACGTGCCCTTCAGAGAGTCAGATGGC 123

Qy   2593 CTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCC 2652
Db    124 CTGGTCTGGGCTGAAGAGTTCAGGTGTAAAGGAGAGGAGCCTGAGCTCTGGTCCTGCCCC 183

Qy   2653 ATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCC 2712
Db    184 AGAGTGCCCTGTCCAGGGGGCACTTGTCACCACAGTGGAGCTGTTCATGTTGTCTGTT-- 241

Qy   2713 CGATATACAGATGTCCGACTTGTG---AATGGCAAATCCCAGTGTGACGGGCAAGTGGAG 2769
Db    242 -------CAGAAGTCCGACTCATGAAAAATGGCTCCTCTCAATGTGAGGGGCAGGTGGAG 294

Qy   2770 ATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCC 2829
Db    295 ATGAATATTTCTGGACGGTGGAGAGCACTCTGTGCCTCCCACTGGAGTCTGGCCAATGCC 354

Qy   2830 CGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATAT 2889
Db    355 AATGTTGTCTGTCGTCAGCTCGGCTGTGGAGTCGCCATCTCCACCCCTAGAGGACCACAC 414

Qy   2890 ATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCA 2949
Db    415 TTTGTGGAAGGAAGTGATCAGATCTCAACAGTCCGATTTCACTGTTCCGGGGTGGAGTCC 474

Qy   2950 CTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACT 3009
Db    475 TTTCTGTGGAGTTGTCCTGTGACTGCTCTGGGTGGTCCTGACTGTTCCCATGGCAACATG 534

Qy   3010 GTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTA 3069
Db    535 GCCTCTGTGATCTGCTCAGGAAACCAGACCCAGGTGCTGCCCCAGTGCAACGACTCCATG 594
```

```
Qy   3070 TCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAA 3129
Db    595 TCTGAAAGAGCAGGCTCTGTAGTCTCAGAAGAGATCGCCCCCTACTGCTCAGATAGCAGA 654

Qy   3130 CGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGAC 3189
Db    655 CAGCTCCGCCTGGTGGACGGGGACAGTCCCTGTGCCGGGAGAGTGGAGATCCTCGACCAG 714

Qy   3190 GGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGT 3249
Db    715 GGCTCCTGGGGCACCATCTGTGATGACGGCTGGGACCTGGATGATGCCCGCGTGGTGTGC 774

Qy   3250 CAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGG 3309
Db    775 AGGCAGTTGGGCTGTGGAGAAGCCCTCAATGCCACAGGGTCTGCTCACTTCGGGGCAGGA 834

Qy   3310 TCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAG 3369
Db    835 TCAGGGCCCATCTGGCTGAACGACTTGAACTGCACAGGAAATGAGTCTCACGTGTGGAAC 894

Qy   3370 TGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATC 3429
Db    895 TGCCCTTCCCGGGGCTGGGGGCGGCACGACTGCAGACACAAGGAGGACGCGGGGGTCATC 954

Qy   3430 TGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGG 3489
Db    955 TGCTCAGAGTTCCTGGCC------CTCAGAATGGTGAGCGAGGACCAGCAGTGTGCTGGG 1008

Qy   3490 AGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACA 3549
Db   1009 TGGCTGGAAGTTTTCTACAATGGGACCTGGGGCAGTGTCTGCCGAAGCCCCATGGAAGAC 1068

Qy   3550 GCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTC 3609
Db   1069 ATCACTGTGTCCATCATCTGCAGTCAGCTTGGCTGTGGGGACAGTGGAAGTCTCAACTCT 1128

Qy   3610 GCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAA 3669
Db   1129 TCTGTTGCTCTTAGAGAAGGTTTAGACCTCGGTGGGTAGATGGAATCCGGTGTCGGAAC 1188

Qy   3670 ACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCA 3729
Db   1189 ACAGATACTTCTCTCTGGCAGTGTCCATCTGACCATTGGAATTATACTTCATGCTCTCCA 1248

Qy   3730 GCAGAAGAGACCTGGATCACATGTGAAGATAGAA------TAAGAGTGCGTGGAGGAGAC 3783
Db   1249 AAGGAGGAAGCCTACATCTCATGTGCAGGCAGCAGACAGCTCCGCCTGGTGGACGGGGGC 1308

Qy   3784 ACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGAT 3843
Db   1309 GGTCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGAT 1368

Qy   3844 GACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCT 3903
Db   1369 GACAGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGGAGCC 1428

Qy   3904 CTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGAC 3963
Db   1429 CTCAATGCCTTGCAATTTGCTGGATTAGGTCAGGGATCAGGGCCCATCTGGCTGGATGAA 1488

Qy   3964 ATGCGGTGCAAAGGAAATGAGTCATTTCTATGGACTGTCACGCCAAACCCTGGGGACAG 4023
Db   1489 CTGAACTGCGGAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCAGGGCTGGGGGCGG 1548

Qy   4024 AGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGA 4068
Db   1549 CACGACTGTGGGCATGCAGAGGATGCTGGAATCGTCTGCTCAGGA 1593

RESULT 12
SSSRP3
LOCUS       SSSRP3                  1665 bp    DNA     linear   MAM 27-AUG-1997
DEFINITION  S.scrofa mRNA for scavenger-receptor protein, 1665bp.
ACCESSION   X99334
VERSION     X99334.1  GI:1480360
KEYWORDS    WC1 gene.
SOURCE      Sus scrofa.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1665)
  AUTHORS   Kanan,J.H., Nayeem,N., Binns,R.M. and Chain,B.M.
  TITLE     Mechanisms for variability in a member of the scavenger-receptor
            cysteine-rich superfamily
  JOURNAL   Immunogenetics 46 (4), 276-282 (1997)
  MEDLINE   97364683
   PUBMED   9218529
REFERENCE   2  (bases 1 to 1665)
  AUTHORS   Kanan,J.H.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1996) J.H.C. Kanan, University College London,
            Immunology, Fourth floor, Windeyer Building, Cleveland Street,
            London WC1 6BT, UK
FEATURES             Location/Qualifiers
     source          1..1665
                     /organism="Sus scrofa"
                     /strain="SLAb/b"
                     /db_xref="taxon:9823"
                     /cell_type="peripheral blood T lymphocyte"
     gene            29..1665
                     /gene="WC1"
     CDS             29..>1665
                     /gene="WC1"
                     /codon_start=1
                     /product="scavenger-receptor protein"
                     /protein_id="CAA67708.1"
                     /db_xref="GI:1480361"
                     /db_xref="SPTREMBL:Q29111"
                     /translation="MALDRHLSLQGLCFLLLIVVGGQALELRLKDGRHRCEGRVELKH
                     QGKWGTVNQAYWGLPGATVVCRQLECGTAVDAPVASYFGRGDGPIWLYAFCEGTESTL
                     TDCRHPNITDYRTKGFSHDWDAGAVCSGFVRLVVGDRPCSGRVEVHSGEDWTPVSDGN
                     FTFPTAHVICAELGCGKAVSVLGQRPLREPGNQVWPEEFRCKGDEPELWSCPRVPCPG
                     GTCHHSGAVQVVCSAYSEVRLMRNGTSQCAGRVEMNISGGWRTLCASHWTMANANVVC
                     HQLGCGVALSTPKGAEGGDQMWKARFHCSGAESFLWNCPVTALGVPDCAHGNTASVIC
                     SGNQTQVLPQCNDSLSEPAGSAASEESTTNCSESRQLRLVDGESYCSRRVEILHQGSW
                     GTICDESWDLDDAHVVCRQLGCGDAISALGSAHFGVGSGRIWLDDVNCTGKESHLWQC
                     PFRGWGRHDCRHKEDAGVICSESLALRMVSEDQECAGWLEVFYNGTWGGVCRSPMEDI
                     TLSMICRQLGCGDSGTLNSSVSLREGSRPRWVDGIQCG"
     sig_peptide     29..94
                     /gene="WC1"
BASE COUNT      344 a    405 c    539 g    377 t
ORIGIN

  Query Match             11.9%;  Score 551.8;  DB 4;  Length 1665;
  Best Local Similarity   61.6%;  Pred. No. 1.4e-126;
  Matches  974;  Conservative    0;  Mismatches  587;  Indels   20;  Gaps    5;

Qy   2086 GATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAAT 2145
Db     98 GCTCTGGAGCTGAGGCTGAAGGATGGAAGGCACCGCTGTGAGGGGAGAGTGGAATTGAAG 157

Qy   2146 GTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTT 2205
Db    158 CACCAAGGAAAATGGGGCACTGTGAATCAGGCTTACTGGGGTTTGCCAGGAGCAACCGTG 217

Qy   2206 GTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACA 2265
Db    218 GTGTGCAGACAGCTGGAATGTGGAACTGCTGTTGATGCTCCCGTGGCATCTTATTTTGGA 277

Qy   2266 GAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAG-GGGAAGCCTCTCT 2324
Db    278 CGAGGGGATGGTCCCATCTGGTTGTATGCTTTTTGTGAAGGGACAGAGTCAACTCTCACT 337

Qy   2325 CTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAG 2384
```

```
Db    338 GACTGTAGGCATCCTAATATTACAGACTACAGGACTAAAGGCTTTTCCCATGATTGGGAT 397

Qy   2385 TTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGG 2444
Db    398 GCTGGAGCAGTCTGCTCAGGATTTGTGC-GCCTGGTTGTAGGGGACAGACCTTGCTCAGG 456

Qy   2445 ACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCT 2504
Db    457 GCGAGTAGAAGTGCATTCTGGAGAAGACTGGACGCCAGTGTCTGATGGGAACTTTACATT 516

Qy   2505 TCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGT 2564
Db    517 CCCCACTGCCCACGTCATCTGTGCAGAGCTGGGATGTGGCAAGGCTGTGTCTGTCTTGGG 576

Qy   2565 GGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGG 2624
Db    577 TCAGAGGCCCCTCAGAGAGCCAGGTAACCAGGTCTGGCCTGAAGAGTTCAGGTGTAAAGG 636

Qy   2625 GAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCA 2684
Db    637 GGACGAGCCTGAGCTCTGGTCCTGCCCCAGAGTGCCCTGTCCAGGAGGCACTTGTCACCA 696

Qy   2685 CAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTG---AATGG 2741
Db    697 CAGTGGAGCTGTTCAAGTTGTCTGTTCAGCATACTCAGAAGTCCGGCTCATGAGAAATGG 756

Qy   2742 CAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTG 2801
Db    757 CACCTCTCAGTGTGCGGGGCGAGTGGAGATGAATATTTCTGGAGGATGGAGAACACTCTG 816

Qy   2802 TGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGAC 2861
Db    817 TGCCTCCCACTGGACCATGGCCAATGCCAATGTTGTGTGTCATCAGCTCGGCTGTGGAGT 876

Qy   2862 TGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACA 2921
Db    877 CGCCCTCTCCACCCCGAAGGGA---------GCAGAGGGAGGTGATCAGATGTGGAAAGC 927

Qy   2922 CAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGG 2981
Db    928 CCGATTTCACTGCTCTGGGGCTGAGTCCTTCTTGTGGAATTGCCCTGTGACTGCCCTGGG 987

Qy   2982 AGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCA 3041
Db    988 TGTTCCTGACTGTGCCCATGGAAACACAGCCTCTGTGATCTGCTCAGGAAACCAGACCCA 1047

Qy   3042 GCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGG 3101
Db   1048 GGTGCTGCCTCAGTGCAATGACTCCCTGTCTGAACCAGCAGGCTCTGCAGCCTCAGAGGA 1107

Qy   3102 CAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTG 3161
Db   1108 AAGCACCACCAACTGCTCAGAGAGCAGACAGCTGCGCCTGGTGGATGGAGAAAGTTATTG 1167

Qy   3162 TGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTG 3221
Db   1168 CTCACGGAGAGTGGAGATCCTTCACCAAGGCTCCTGGGGCACCATCTGTGATGAAAGCTG 1227

Qy   3222 GGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGC 3281
Db   1228 GGACCTGGACGATGCGCACGTGGTGTGCAGGCAGCTGGGCTGTGGAGATGCCATCAGTGC 1287

Qy   3282 CACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTG 3341
Db   1288 ACTGGGGTCTGCTCACTTTGGGGTGGGCTCAGGACGCATCTGGCTGGATGACGTGAACTG 1347

Qy   3342 CACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTG 3401
Db   1348 TACAGGAAAGGAGTCCCACCTATGGCAGTGCCCTTTCAGGGGCTGGGGGCGGCACGACTG 1407

Qy   3402 CAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAG 3461
Db   1408 CAGACACAAGGAGGACGCAGGGGTCATCTGCTCAGAGTCCCTGGCCCTCAGG------AT 1461

Qy   3462 TGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGG 3521
Db   1462 GGTGAGCGAGGACCAGGAGTGTGCCGGGTGGCTGGAAGTTTTCTACAACGGGACCTGGGG 1521

Qy   3522 CAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGG 3581
Db   1522 CGGTGTCTGCCGCAGCCCCATGGAAGACATCACCTTGTCCATGATCTGCAGGCAGCTTGG 1581

Qy   3582 CTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCAT 3641
Db   1582 CTGTGGGGACAGTGGAACCCTCAACTCTTCCGTTTCTCTTAGGGAAGGTTCTAGACCCCG 1641

Qy   3642 GTGGGTGGATGACATTCAGTG 3662
Db   1642 GTGGGTCGATGGAATCCAGTG 1662
```

RESULT 13
AC131207/c

```
LOCUS       AC131207               202252 bp    DNA     linear   HTG 18-AUG-2002
DEFINITION  Homo sapiens clone RP11-326A16, *** SEQUENCING IN PROGRESS ***, 10
            unordered pieces.
ACCESSION   AC131207
VERSION     AC131207.1  GI:22296912
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 202252)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 202252)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HELB
                Center clone name: RP11-326A16
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 194619 bases at least Q40
                Consensus quality: 196802 bases at least Q30
                Consensus quality: 198233 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2422: contig of 2422 bp in length
            *     2423     2522: gap of unknown length
            *     2523     7722: contig of 5200 bp in length
            *     7723     7822: gap of unknown length
            *     7823    13397: contig of 5575 bp in length
            *    13398    13497: gap of unknown length
            *    13498    22395: contig of 8898 bp in length
            *    22396    22495: gap of unknown length
            *    22496    39888: contig of 17393 bp in length
            *    39889    39988: gap of unknown length
            *    39989    58697: contig of 18709 bp in length
            *    58698    58797: gap of unknown length
            *    58798    81609: contig of 22812 bp in length
            *    81610    81709: gap of unknown length
            *    81710   106273: contig of 24564 bp in length
            *   106274   106373: gap of unknown length
            *   106374   143244: contig of 36871 bp in length
            *   143245   143344: gap of unknown length
            *   143345   202252: contig of 58908 bp in length.
FEATURES             Location/Qualifiers
     source          1..202252
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-326A16"
BASE COUNT    60981 a  37684 c  37424 g  62900 t   3263 others
ORIGIN

  Query Match             11.2%;  Score 517.8;  DB 2;  Length 202252;
  Best Local Similarity   85.7%;  Pred. No. 5.6e-118;
  Matches  633;  Conservative    0;  Mismatches    2;  Indels  104;  Gaps    1;

Qy    2397 AGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGT 2456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72800 AGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGT 72741

Qy    2457 GAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAA 2516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72740 GAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAA 72681

Qy    2517 TGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTT 2576
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   72680 TGTGCTGTGCAGAGAATTAAACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTT 72621

Qy    2577 TGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCA 2636
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72620 TGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCA 72561

Qy    2637 CCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGT 2696
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72560 CCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGT 72501

Qy    2697 TGGAGTTGTCTGTTCCC------------------------------------------- 2713
           |||||||||||||||||
Db   72500 TGGAGTTGTCTGTTCCCGTGAGATTTTTAAATAATCTTAATTGGCTGGGGCTAGAGATGG 72441

Qy    2714 ------------------------------------------------------------ 2713

Db   72440 GGAGTTTGGGGCTGGAGCTGATAAGTCATGATGTCTCTTCTATGGCCTTGTCTCCTCCTA 72381

Qy    2714 -GATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATC 2772
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72380 GGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATC 72321

Qy    2773 AACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGT 2832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72320 AACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGT 72261

Qy    2833 GTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATT 2892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72260 GTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATT 72201

Qy    2893 GGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTT 2952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72200 GGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTT 72141

Qy    2953 CTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTC 3012
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72140 CTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTC 72081

Qy    3013 TCTGTGATCTGCACAGGAA 3031
           ||||||||||||||||| |
Db   72080 TCTGTGATCTGCACAGGTA 72062


RESULT 14
AF400122
LOCUS       AF400122                2165 bp    mRNA    linear   MAM 21-AUG-2001
DEFINITION  Ovis aries EP4a mRNA, partial cds.
ACCESSION   AF400122
VERSION     AF400122.1  GI:15217046
KEYWORDS    .
SOURCE      Ovis aries.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 2165)
  AUTHORS   Gill,J.W. and Walker,I.D.
  TITLE     Novel T19 transcripts from ovine CD4+ T cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2165)
  AUTHORS   Gill,J.W. and Walker,I.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2001) Veterinary Science, University of
            Melbourne, Cnr Park Drive and Flemington Rd, Parkville, VIC 3052,
            Australia
FEATURES             Location/Qualifiers
     source          1..2165
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
                     /cell_type="CD4+ T cell"
                     /tissue_type="efferent prescapular lymph"
     CDS             <1..1263
                     /note="T19 multigene family member"
```

```
                    /codon_start=1
                    /product="EP4a"
                    /protein_id="AAK92475.1"
                    /db_xref="GI:15217047"
                    /translation="RRSLLHCPISLSSDSRQIRLVDGGGSCVGRVEILDQGSWGTICD
                    DGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDDLNCAGNESHVWRCPSRGW
                    GRHDCRHKEDAGVICSEYLALRMVSEDQECAGWLEVFYNGTWGSVCRSPMDDVTVSII
                    CSQLGCGDSGSLNTSVGLREGSRPRWIDLIQCRKTDTSLWQCPSGPWKFSSCSPKEEA
                    YISCAGRRPKSCPAAAPCTDREKLRLRGGDSECSGRVEVWHSGSWGTVCDDSWSLAEA
                    EVVCQQLGCGQALEAVRAGTFGPGNGSIWLDEVRCRGRESSLWDCAAEPWGQSDCKHE
                    EDAGVRCSGVRKTLPPTPAGTVNIHGLERILRSGDLFCGSLTAQERKEPEEPWLVGRS
                    PGRYVGELLCLGSWGHTFLMSGVQCCRI"
BASE COUNT      420 a    566 c    707 g    470 t     2 others
ORIGIN

  Query Match             10.3%;  Score 476.4;  DB 4;  Length 2165;
  Best Local Similarity   69.5%;  Pred. No. 1e-107;
  Matches  701;  Conservative    1;  Mismatches  252;  Indels   54;  Gaps    2;

Qy       3108 TTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGG 3167
              ||   ||| || |||    || || | ||||||| ||||| |||| | |     ||| | |||
Db         27 TTCTCTCTCCTCAGACAGCAGACAGATCCGCCTGGTGGACGGGGGCGGTTCCTGCGTCGG 86

Qy       3168 GAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCT 3227
              |||||| ||||||  | || |  ||||  |||||||||||||||||||||||||||||||
Db         87 GAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGACGGCTGGGACCT 146

Qy       3228 GAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGT 3287
              | |||||||| |||||||||        |||||||||||||||  ||| ||||||||||||
Db        147 GGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGCCCTCAATGCCACGGG 206

Qy       3288 CTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGG 3347
              |||||||||||  ||||   || ||||||||||||||||||||| ||||||||||||| | ||
Db        207 GTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGCTGGACGACCTGAACTGCGCCGG 266

Qy       3348 AACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGCAGCACGACTGCAGGCA 3407
              ||  ||||| ||| |||||  ||||||||||||| |||||||| ||||||||||||| ||
Db        267 AAATGAGTCTCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCGGCACGACTGCAGACA 326

Qy       3408 CAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAAC 3467
              |||||||||||| ||||||||||||||||| | |    |||  | |||       |   |   |
Db        327 CAAGGAGGACGCGGGGGTCATCTGCTCAGAGTACCTGGCCCTCAGG------ATGGTGAG 380

Qy       3468 TGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGT 3527
               ||      ||    || |||||| |  ||||||| ||||| |||||||||||||||| ||
Db        381 CGAGGACCAGGAGTGCGCTGGGTGGCTGGAAGTTTTCTACAACGGGACCTGGGGCAGTGT 440

Qy       3528 CGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGG 3587
              | || |  ||     |||            | ||   |       |||   | ||||| ||||| ||||||||
Db        441 CTGCCGCAGCCCCATGGATGATGTCACCGTGTCCATCATCTGCAGTCAGCTTGGCTGTGG 500

Qy       3588 GGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGT 3647
              |||  |  ||||  | |||  |         |  | | |||  |||  |   |   ||||  |
Db        501 GGACAGTGGAAGTCTCAACACTTCTGTTGGTCTCAGGGAAGGTTCTAGACCCCGGTGGAT 560

Qy       3648 GGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATG 3707
              |||    ||||||||||   |||||   ||| |||    | ||||||||  |   |||| ||| ||
Db        561 AGATTTAATTCAGTGTCGGAAAACTGATACCTCTCTCTGGCAGTGTCCTTCTGGCCCGTG 620

Qy       3708 GGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAA- 3766
              | |      ||    || |     |||     || ||   |||   |||  | |||| ||   |||| |
Db        621 GAAATTCAGTTCCTGCTCTCCAAAGGAGGAAGCCTACATCTCCTGTGCAGGAAGAAGACC 680

Qy       3767 ---------------------------------------------GAGTGCGTGGAGG 3779
                                                           |  |   | |||||
Db        681 CAAGAGCTGTCCAGCTGCTGCCCCCTGCACAGACAGAGAGAAGCTCCGGCTCAGGGGAGG 740

Qy       3780 AGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTG 3839
              ||||| ||||||||| ||| |   ||| | |||||||  ||||||||||||  |||||
Db        741 AGACAGCGAGTGCTCAGGGCGGGTGGAGGTCTGGCACAGCGGCTCCTGGGGCACCGTGTG 800

Qy       3840 TGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTC 3899
               |||||||||||   |||||| ||||| ||  ||||||||||||||||||||||||||||
Db        801 CGATGACTCCTGGAGCCTGGCAGAGGCCGAGGTGGTGTGTCAGCAGCTGGGCTGTGGCCA 860

Qy       3900 TGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGA 3959
                || ||||   ||| || |||   |    | || ||||   |||| |||  | ||||||| ||||
Db        861 GGCCCTGGAAGCCGTGCGGGCTGGGACATTCGGCCCTGGAAATGGGAGCATCTGGCTGGA 920

Qy       3960 TGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGG 4019
                ||   |||||||||   ||         |||||  |   ||  |||||||||      |||  |  ||||||||
Db        921 CGAGGTGCGGTGCAGGGGCCGGGAGTCCTCCCTGTGGGACTGTGCTGCCGAGCCCTGGGG 980

Qy       4020 ACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGG 4067
                ||||| |||||    :||| |||| || ||||| ||||||||||||||
Db        981 GCAGAGCGACTGCAARCACGAGGAGGACGCTGGTGTGAGGTGCTCTGG 1028


RESULT 15
SSSRP2
LOCUS       SSSRP2                  1316 bp    DNA     linear   MAM 27-AUG-1997
DEFINITION  S.scrofa mRNA for scavenger-receptor protein, 1316bp.
ACCESSION   X99333
VERSION     X99333.1  GI:1480358
KEYWORDS    WC1 gene.
SOURCE      Sus scrofa.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1316)
  AUTHORS   Kanan,J.H., Nayeem,N., Binns,R.M. and Chain,B.M.
  TITLE     Mechanisms for variability in a member of the scavenger-receptor
            cysteine-rich superfamily
  JOURNAL   Immunogenetics 46 (4), 276-282 (1997)
  MEDLINE   97364683
   PUBMED   9218529
REFERENCE   2  (bases 1 to 1316)
  AUTHORS   Kanan,J.H.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1996) J.H.C.-Kanan, University College London,
            Immunology, Fourth floor, Windeyer Building, Cleveland Street,
            London WC1 6BT, UK
FEATURES             Location/Qualifiers
     source          1..1316
                     /organism="Sus scrofa"
                     /strain="SLAb/b"
                     /db_xref="taxon:9823"
                     /cell_type="peripheral blood T lymphocyte"
     intron          <1..126
                     /number=1
     gene            join(127..219,301..1316)
                     /gene="WC1"
     CDS             join(<127..219,301..>1316)
                     /gene="WC1"
                     /codon_start=3
                     /product="scavenger-receptor protein"
                     /protein_id="CAA67707.1"
                     /db_xref="GI:1480359"
                     /db_xref="SPTREMBL:Q29110"
                     /translation="NQTQVLPQCNDSLSEPAGSAASEEGTANCSDSRQLRLVDGGSYC
                     SGSVEILHQGSWGTVCDDSWDLDDAHVVCRQLGCGKTLSALGSAHFGAGSGRIWLDDV
                     NCTGKESHLWQCPSRGWGQHNCRHKEDAGVICSESLALRMVSEDQECAGWLEVFYNGT
                     WGGVCRSPMEDITLSIICRQLGCGDSGTLNSSVSLREGSRPRWVDRIQCRKTDSSLWE
                     CPSDPWKYSSCSSKEEAYVSCAGRRPKSCPTTAPCTDKEKLRLRGGDSVCSGRVEVWH
                     GGSWGTVCDDSWSLAEAEVVCQQLGCGHALEALRDAAFGPGSGSIWLDEVQCRGRESS
                     LWDCAAGPWGQSDCKHEEDAGVRCSGARPGSNPIP"
     exon            127..219
                     /gene="WC1"
                     /number=2
     intron          220..300
                     /gene="WC1"
                     /number=2
     exon            301..>1316
                     /gene="WC1"
                     /number=3
BASE COUNT      264 a    345 c    422 g    285 t
```

```
ORIGIN

  Query Match             10.0%;  Score 460.6;  DB 4;  Length 1316;
  Best Local Similarity   68.8%;  Pred. No. 8.7e-104;
  Matches  690;  Conservative    0;  Mismatches  259;  Indels   54;  Gaps    2;

Qy       3113 TCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAG 3172
Db        292 TCTATTCAGACAGCAGACAGCTGCGCCTGGTGGATGGAGGGAGTTACTGCTCAGGGAGCG 351

Qy       3173 TAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCG 3232
Db        352 TGGAGATCCTTCACCAGGGCTCCTGGGGCACCGTCTGTGATGACAGCTGGGACCTGGACG 411

Qy       3233 ATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTG 3292
Db        412 ATGCGCATGTGGTGTGCAGGCAGCTAGGCTGTGGAAAAACCCTCAGTGCACTGGGGTCTG 471

Qy       3293 CTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGG 3352
Db        472 CTCACTTTGGGGCGGGCTCAGGACGCATCTGGCTGGATGACGTGAACTGTACAGGAAAGG 531

Qy       3353 AGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGG 3412
Db        532 AGTCCCACCTATGGCAGTGCCCTTCCCGGGGCTGGGGGCAGCACAACTGCAGACACAAGG 591

Qy       3413 AGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAA 3472
Db        592 AGGACGCAGGGGTCATCTGCTCAGAGTCCCTGGCCCTCAGG------ATGGTGAGCGAGG 645

Qy       3473 CAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCA 3532
Db        646 ACCAGGAGTGTGCCGGGTGGCTGGAAGTTTTCTACAACGGGACCTGGGGCGGTGTCTGCC 705

Qy       3533 GGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGA 3592
Db        706 GCAGCCCCATGGAAGACATCACCTTGTCCATCATCTGCAGGCAGCTTGGCTGTGGGGACA 765

Qy       3593 ATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATG 3652
Db        766 GTGGAACCCTCAACTCTTCTGTTTCTCTTAGGGAAGGTTCTAGACCCCGGTGGGTGGATA 825

Qy       3653 ACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGC 3712
Db        826 GAATCCAGTGTCGGAAAACTGATAGCTCTCTCTGGGAGTGTCCTTCTGACCCTTGGAAAT 885

Qy       3713 GAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAA------ 3766
Db        886 ATAGTTCATGTTCTTCAAAGGAGGAAGCCTATGTCTCCTGTGCAGGAAGGAGACCCAAGA 945

Qy       3767 ----------------------------------------GAGTGCGTGGAGGAGACA 3784
Db        946 GCTGTCCAACTACTGCCCCCTGCACAGACAAAGAGAAGCTGCGACTCAGGGGAGGAGACA 1005

Qy       3785 CCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATG 3844
Db       1006 GCGTGTGCTCAGGGCGAGTGGAGGTGTGGCACGGCGGCTCCTGGGGCACTGTGTGTGATG 1065

Qy       3845 ACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTC 3904
Db       1066 ACTCCTGGAGCCTGGCAGAGGCCGAGGTGGTGTGTCAGCAGCTGGGCTGTGGCCATGCCC 1125

Qy       3905 TGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACA 3964
Db       1126 TGGAAGCCCTGCGGGATGCGGCATTTGGCCCTGGAAGTGGAGCATCTGGCTGGACGAGG 1185

Qy       3965 TGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGA 4024
Db       1186 TGCAGTGCAGGGGCAGGGAGTCCTCCCTGTGGGACTGTGCTGCGGGGCCCTGGGGGCAGA 1245

Qy       4025 GTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGG 4067
Db       1246 GTGACTGCAAGCATGAGGAGGACGCTGGAGTGAGATGCTCTGG 1288
```

Search completed: May 12, 2003, 05:29:51
Job time : 8566 secs

```
OM nucleic - nucleic search, using sw model

Run on:          May 12, 2003, 02:53:40 ; Search time 3980 Seconds
                                          (without alignments)
                                          18832.340 Million cell updates/sec

Title:           US-09-759-130B-379
Perfect score:   4628
Sequence:        1 gcggccgctcgcgatctaga..........aaacggacgcgtgggtcgac 4628

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	958.2	20.7	999	9	AL529439	AL529439 AL529439
	2	578.8	12.5	582	12	BG570972	BG570972 602591546
c	3	568.2	12.3	581	14	BQ776070	BQ776070 UI-H-FH0-
c	4	512.6	11.1	554	9	AL529438	AL529438 AL529438
	5	491.4	10.6	940	9	AL550994	AL550994 AL550994
c	6	489.4	10.6	491	9	AI095706	AI095706 qb24e01.x
c	7	453.8	9.8	457	9	AI738432	AI738432 wi23g06.x
c	8	447.4	9.7	449	9	AI609329	AI609329 tw30b02.x
c	9	446.2	9.6	451	10	AW082879	AW082879 xb71g11.x
	10	416.8	9.0	799	13	BI185454	BI185454 UNL-P-FN-
c	11	414.4	9.0	429	9	AI278132	AI278132 ql77e02.x
	12	408.4	8.8	718	13	BI183796	BI183796 UNL-P-FN-
c	13	407.8	8.8	411	9	AI090561	AI090561 qa64g07.x
	14	401.4	8.7	483	12	BF132072	BF132072 601821079
c	15	392.8	8.5	606	14	BQ024006	BQ024006 UI-1-BB1p
c	16	330.8	8.2	392	14	W42792	W42792 zc25b03.s1
c	17	359.4	8.0	371	14	N69291	N69291 za25d06.s1
	18	359.4	8.0	385	14	W42886	W42886 zc25b03.r1
	19	366	7.9	822	12	BF792807	BF792807 602253274
	20	363.8	7.9	1100	9	AL540905	AL540905 AL540905
	21	363.4	7.9	1108	9	AL544534	AL544534 AL544534
	22	354	7.6	930	13	BI838644	BI838644 603086185
	23	324.2	7.0	705	17	B16308	B16308 34707.TP CI
c	24	320	6.9	320	9	AA582301	AA582301 nn49d01.s
	25	310	6.7	723	13	BI836459	BI836459 603082833
c	26	308.8	6.7	1153	9	AL540904	AL540904 AL540904
	27	306.8	6.6	310	9	AL706449	AL706449 DKFZp686P
	28	304.8	6.6	484	14	N94111	N94111 za25d06.r1
	29	301.8	6.5	644	13	BI009838	BI009838 MR4-RT004
	30	299.2	6.5	653	10	BE502724	BE502724 hz78d12.x
c	31	296.8	6.4	616	13	BI009795	BI009795 MR4-RT004
c	32	294.8	6.4	310	9	AI472927	AI472927 tj70h06.x
	33	291.6	6.3	585	14	BQ327934	BQ327934 MR4-RT004
c	34	290.4	6.3	408	9	AI288981	AI288981 ql86e12.x
c	35	289.6	6.3	711	14	BQ378694	BQ378694 MR3-UT009
c	36	277.4	6.0	583	14	BQ368412	BQ368412 PM3-GN051
c	37	276.8	6.0	689	12	BF096204	BF096204 MR3-UT009
	38	275.2	5.9	569	10	AW631641	AW631641 90867 MAR
c	39	275.2	5.9	596	13	BI009802	BI009802 MR4-RT004
	40	269.8	5.8	574	14	BQ327932	BQ327932 MR4-RT004
	41	268.2	5.8	452	17	AQ672871	AQ672871 HS_5502_B
c	42	267	5.8	833	9	AL574811	AL574811 AL574811
	43	266.4	5.8	574	13	BI009286	BI009286 MR4-RT004
c	44	265.4	5.7	560	13	BI009799	BI009799 MR4-RT004
	45	264.4	5.7	583	14	BQ327909	BQ327909 MR4-RT004

ALIGNMENTS

```
RESULT 1
AL529439
LOCUS       AL529439                 999 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL529439 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD006YO05 5
            prime, mRNA sequence.
ACCESSION   AL529439
VERSION     AL529439.1  GI:12792932
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 999)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..999
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DD006YO05"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
```

```
                    /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                    was primed with a NotI-oligo(dT) primer. Five prime end
                    enriched, double-stranded cDNA was digested with Not I and
                    cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                    vector.  Library was normalized. Library was constructed
                    by Life Technologies.  Contact : Feng Liang Life
                    Technologies, a division of Invitrogen 9800 Medical Center
                    Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                    8371 Email : fliang@lifetech.com URL :
                    http://fulllength.invitrogen.com"
BASE COUNT      220 a    236 c    318 g    224 t      1 others
ORIGIN

  Query Match             20.7%;  Score 958.2;  DB 9;  Length 999;
  Best Local Similarity   99.3%;  Pred. No. 6.8e-154;
  Matches  982;  Conservative    1;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       3205 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGT 3264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGT 60

Qy       3265 GGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGG 3324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGG 120

Qy       3325 CTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGC 3384
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 CTGGATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGC 180

Qy       3385 TGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACA 3444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACA 240

Qy       3445 GCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTC 3504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTC 300

Qy       3505 TATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATT 3564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATT 360

Qy       3565 GTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAG 3624
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        361 GTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTAGTCAGCCTCGCCCCTTTATCTAAG 420

Qy       3625 ACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATA 3684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATA 480

Qy       3685 TGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGG 3744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGG 540

Qy       3745 ATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTG 3804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTG 600

Qy       3805 GAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAG 3864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAG 660

Qy       3865 GCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCT 3924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCT 720

Qy       3925 TCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAG 3984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAG 780

Qy       3985 TCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAA 4044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAA 840

Qy       4045 GATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCAT 4104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCAT 900

Qy       4105 TTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTT 4164
              |||||||||||||||||||||||||||||||| |||||||||||||||| || |||||||
Db        901 TTAGCACTTATTTTATCCAGTATCTTTGGGCT-CTTCTCCTGGTTCTGTATA-TCTATTT 958

Qy       4165 CTCACGTGGTGCCGAGTTCAGAAACAAAA 4193
              |||||||||||||||||||||||: ||||
Db        959 CTCACGTGGTGCCGAGTTCAGAAMAAAAA 987


RESULT 2
BG570972
LOCUS       BG570972                 582 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602591546F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4714115 5',
            mRNA sequence.
ACCESSION   BG570972
VERSION     BG570972.1  GI:13578625
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1558  row: k  column: 12
            High quality sequence stop: 576.
FEATURES             Location/Qualifiers
     source          1..582
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4714115"
                     /clone_lib="NIH_MGC_79"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      126 a    140 c    198 g    118 t
ORIGIN

  Query Match             12.5%;  Score 578.8;  DB 12;  Length 582;
  Best Local Similarity   99.7%;  Pred. No. 4.6e-89;
  Matches  580;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3351 GGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAA 60

Qy       3411 GGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGA 3470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGA 120
```

```
Qy    3471 AACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGG 3530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGG 180

Qy    3531 CAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGA 3590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGA 240

Qy    3591 GAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGA 3650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGA 300

Qy    3651 TGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGA 3710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGA 360

Qy    3711 GCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGT 3770
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     361 GCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATCGAATAAGAGT 420

Qy    3771 GCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGG 3830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGG 480

Qy    3831 CACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGG 3890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGG 540

Qy    3891 CTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGG 3932
           |||||||||||||||||||||||||||||||||||||| |||
Db     541 CTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTATGG 582


RESULT 3
BQ776070/c
LOCUS       BQ776070                 581 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  UI-H-FH0-bcl-f-19-0-UI.s1 NCI_CGAP_FH0 Homo sapiens cDNA clone
            UI-H-FH0-bcl-f-19-0-UI 3', mRNA sequence.
ACCESSION   BQ776070
VERSION     BQ776070.1  GI:21984546
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 1-39, >AT_rich#Low_complexity
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..581
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FH0-bcl-f-19-0-UI"
                     /clone_lib="NCI_CGAP_FH0"
                     /tissue_type="Human Chondrosarcoma Cell Line"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_FH0 is a cDNA library containing the following
                     tissue(s): Human Grade 1 Chondrosarcoma Cell Line The
                     library was constructed according to Bonaldo, Lennon and
                     Soares, Genome Research, 6:791-806, 1996. First strand
                     cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site. Double stranded cDNA was ligated
                     to an EcoR I adaptor, digested with Not I, and cloned
                     directionally into pT7T3-Pac vector. The oligonucleotide
                     used to prime the synthesis of first-strand cDNA contains
                     a library tag sequence that is located between the Not I
                     site and the (dT)18 tail. The sequence tag for this
                     library is AGAATCCGGC. The cell line was provided by Dr
                     James Martin from University of Iowa
                     TAG_LIB=UI-H-FH0
                     TAG_TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
                     Chondrosarcoma
                     TAG_SEQ=AGAATCCGGC"
BASE COUNT      144 a    125 c    123 g    189 t
ORIGIN

  Query Match             12.3%;  Score 568.2;  DB 14;  Length 581;
  Best Local Similarity   98.6%;  Pred. No. 2.9e-87;
  Matches  573;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    4008 CAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGG 4067
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     581 CAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGG 522

Qy    4068 ACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTAT 4127
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     521 ACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTAT 462

Qy    4128 CTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAA 4187
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     461 CTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAA 402

Qy    4188 ACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTT 4247
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 ACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTT 342

Qy    4248 ATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGA 4307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     341 ATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGA 282

Qy    4308 TGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCC 4367
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     281 TGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCC 222

Qy    4368 TGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 4427
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     221 TGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 162

Qy    4428 ATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGA 4487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     161 ATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGA 102

Qy    4488 AAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGG 4547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101 AAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGG 42

Qy    4548 TGATAATAAAAATTGAATTATGTATATCACTGTTAAAAAAA 4588
           |||||||||||||||||||||||| |  |     |||||||
Db      41 TGATAATAAAAATTGAATTATGTAAAAAAAAAAAAAAAAAA 1


RESULT 4
AL529438/c
LOCUS       AL529438                 554 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL529438 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD006YO05 3
            prime, mRNA sequence.
ACCESSION   AL529438
VERSION     AL529438.1  GI:12792931
KEYWORDS    EST.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DD006YO05"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies. Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      132 a    166 c    127 g    110 t     19 others
ORIGIN

  Query Match             11.1%;  Score 512.6;  DB 9;  Length 554;
  Best Local Similarity   94.8%;  Pred. No. 9.1e-78;
  Matches  525;  Conservative    9;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       3592 AATGGAGTTGTCAGCCTCGCCCCTTTA-TCTAAGACAGGCTCTGGTTTCATGTGGGTGGA 3650

Db        554 AATGGAGTTATCAGCCTCGCCCCTTTATTCTAAGACAGGSTCTGGTTTCATGTGGGTGGA 495

Qy       3651 TGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGA 3710

Db        494 TRACATTCARTGTCCTRAAACGCATGTCTCCATATGGCAGTGCCTSTCTGCCCCATGGGA 435

Qy       3711 GCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGT 3770

Db        434 GCGAAGASTCTCCAGCCCVGCAGAAGAGACCTGGCTMACATGTGAAGATAGAATAAGAGT 375

Qy       3771 GCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGG 3830

Db        374 GCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGG 315

Qy       3831 CACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGG 3890

Db        314 CACAGTGTGTGATVACTCCTGGGACCTGGCCGAGGCGGNAGTGGTGTNTCAGCAGCTGGG 255

Qy       3891 CTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCAT 3950

Db        254 CTGTGGCTCTGCTCTNGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGGACTGGAACCNT 195

Qy       3951 CTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAA 4010

Db        194 CTGGTTGGNTTACATGCGGTGCAAAGGAANTNAGTCATTTCTATGGGACTGTCACGCCAA 135

Qy       4011 ACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACA 4070

Db        134 ACCCTGGGGACAGAGTGACTGTGGACACAAGGGAGANGCTGGCGTTAGGTGCTCTGGACA 75

Qy       4071 GTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTT 4130

Db         74 GTCGCTGNAATCACTGAATGCCTCCTCAGGTCATTTAGGACTTATTTTATCCAGTGTCTT 15

Qy       4131 TGGGCTCCTTCTCC 4144
              ||||||||:|||||
Db         14 TGGGCTCCKTCTCC 1


RESULT 5
AL550994
LOCUS       AL550994                 940 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL550994 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI066YE07 5
            prime, mRNA sequence.
ACCESSION   AL550994
VERSION     AL550994.1  GI:12888516
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 940)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI066YE07"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      228 a    212 c    279 g    218 t      3 others
ORIGIN

  Query Match             10.6%;  Score 491.4;  DB 9;  Length 940;
  Best Local Similarity   71.3%;  Pred. No. 3e-74;
  Matches  645;  Conservative    2;  Mismatches  258;  Indels    0;  Gaps    0;

Qy       2599 TGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTT 2658

Db          9 TGGGCTGAAGAATTCCAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCYCAGTAGCA 68

Qy       2659 CAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATAT 2718

Db         69 CCCCGCCCAGAAGGAACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATAC 128

Qy       2719 ACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTG 2778

Db        129 ACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACG 188

Qy       2779 CTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTA 2838

Db        189 CTTGGTGCCTGGGGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTT 248

Qy       2839 TGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAA 2898

Db        249 TGCCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAA 308

Qy       2899 AGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGAT 2958

Db        309 GGAAATGGTCAGATCTGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGA 368
```

```
Qy   2959 AACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTG 3018
            | ||||   | || | ||| || ||  |      |||        |    |     | ||||||
Db    369 GATTGTCCTGTAACTGCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTA 428

Qy   3019 ATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCA 3078
          |||||| |||||| || | ||||  ||||||    | |||     ||         |   | ||||
Db    429 ATCTGCTCAGGAAACCAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCA 488

Qy   3079 TATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGC 3138
                | || |  ||||||| |  |||||| ||  |||| |||||          |  || |||
Db    489 ACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGC 548

Qy   3139 CTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGG 3198
           || ||  |||| |     | ||||||| |||||||||||||||||||| || |||| ||||
Db    549 CTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGG 608

Qy   3199 GGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTG 3258
          |||||||||||||||||| ||||||||||||| |||||||||||||| ||    | |||||
Db    609 GGCACCATCTGTGATGACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTG 668

Qy   3259 GGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCC 3318
          ||||||||| |||| | |||||||| |    |||||||| |||||||| ||   ||||||||
Db    669 GGCTGTGGAGAGGCCATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCC 728

Qy   3319 ATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCC 3378
          |||||||||||||   |||| ||||   ||||   || |||| | | |||||||||| |||
Db    729 ATCTGGCTGGATGAGATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCA 788

Qy   3379 CGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAA 3438
           | |||||||||||||     | ||||||||||||||||| || || || |||||||||||
Db    789 CACGGCTGGGGGCAGCAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAA 848

Qy   3439 TTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAA 3498
          ||||     |   |:|| ||    ||||||| |      | ||||  |||||| ||| |  |||||
Db    849 TTCATGTCTCTKAGACTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGSGTCTGGAA 908

Qy   3499 GTCTT 3503
          || ||
Db    909 GTTTT 913
```

RESULT 6
AI095706/c
LOCUS   AI095706   491 bp   mRNA   linear   EST 05-OCT-1998
DEFINITION qb24e01.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
            IMAGE:1697208 3', mRNA sequence.
ACCESSION   AI095706
VERSION   AI095706.1 GI:3434682
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 491)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 857   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 446.
FEATURES   Location/Qualifiers
     source   1. .491
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1697208"
                     /clone_lib="Soares_pregnant_uterus_NbHPU"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by M. Fatima Bonaldo."
BASE COUNT   133 a   98 c   107 g   153 t
ORIGIN

```
  Query Match             10.6%;  Score 489.4;  DB 9;  Length 491;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-74;
  Matches  490;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   4094 CCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGT 4153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 CCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGT 432

Qy   4154 TTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTT 4213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 TTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTT 372

Qy   4214 CAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCA 4273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 CAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCA 312

Qy   4274 AGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAG 4333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 AGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAG 252

Qy   4334 ATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTT 4393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 ATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTT 192

Qy   4394 AGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTA 4453
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    191 AGACTTCCAGGGCTCACCAGCTCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTA 132

Qy   4454 AATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGA 4513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 AATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGA 72

Qy   4514 ACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATA 4573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71 ACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATA 12

Qy   4574 TCACTGTTAAA 4584
          |||||||||||
Db     11 TCACTGTTAAA 1
```

RESULT 7
AI738432/c
LOCUS   AI738432   457 bp   mRNA   linear   EST 21-DEC-1999
DEFINITION wi23g06.x1 NCI_CGAP_Co16 Homo sapiens cDNA clone IMAGE:2391130 3',
            mRNA sequence.
ACCESSION   AI738432
VERSION   AI738432.1 GI:5100413
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 457)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.

```
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 600   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2391130"
                     /clone_lib="NCI_CGAP_Co16"
                     /tissue_type="colon tumor, RER+"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Co10 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1057416-1061255, and 1144584-1145351).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT     119 a     94 c    100 g    144 t
ORIGIN

  Query Match              9.8%;  Score 453.8;  DB 9;  Length 457;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-67;
  Matches  455;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    4128 CTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAA 4187
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 CTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAA 398

Qy    4188 ACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTT 4247
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 ACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTT 338

Qy    4248 ATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGA 4307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 ATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGA 278

Qy    4308 TGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCC 4367
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 TGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCC 218

Qy    4368 TGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 4427
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db     217 TGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGCTCAACCTCTAAAT 158

Qy    4428 ATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGA 4487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157 ATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGA 98

Qy    4488 AAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGG 4547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97 AAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGG 38

Qy    4548 TGATAATAAAAATTGAATTATGTATATCACTGTTAAA 4584
           ||||||||||||||||||||||||||||| |||||||
Db      37 TGATAATAAAAATTGAATTATGTATATCCCTGTTAAA 1


RESULT 8
AI609329/c
LOCUS       AI609329               449 bp    mRNA    linear   EST 14-MAY-1999
DEFINITION  tw30b02.x1 NCI_CGAP_Ov35 Homo sapiens cDNA clone IMAGE:2261163 3',
            mRNA sequence.
ACCESSION   AI609329
VERSION     AI609329.1  GI:4618496
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 631   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 415
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2261163"
                     /clone_lib="NCI_CGAP_Ov35"
                     /tissue_type="tumor, 5 pooled (see description)"
                     /lab_host="DH10B"
                     /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; This library represents the normalized
                     version of NCI_CGAP_Ov23. Cloned unidirectionally.
                     Primer: Oligo dT.  Average insert size 0.86 kb. Tumor
                     types include: mixed Mullerian tumor, papillary serous,
                     clear cell, spindle cell.  All are primary tumors,
                     metastasis positive. Constructed by Life Technologies."
BASE COUNT     116 a     94 c     97 g    142 t
ORIGIN

  Query Match              9.7%;  Score 447.4;  DB 9;  Length 449;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-66;
  Matches  448;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    4133 GGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAA 4192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     449 GGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAA 390

Qy    4193 AACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCC 4252
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     389 AACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCC 330

Qy    4253 ATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACA 4312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329 ATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACA 270

Qy    4313 CCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCT 4372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     269 CCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCT 210

Qy    4373 CTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 4432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     209 CTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 150

Qy    4433 TGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACT 4492
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     149 TGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTGTGGAAAACT 90

Qy    4493 TGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 4552
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        89 TGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 30

Qy      4553 ATAAAAATTGAATTATGTATATCACTGTT 4581
             |||||||||||||||||||||||||||||
Db        29 ATAAAAATTGAATTATGTATATCACTGTT 1


RESULT 9
AW082879/c
LOCUS       AW082879                 451 bp    mRNA    linear   EST 14-OCT-1999
DEFINITION  xb71g11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2581796 3', mRNA sequence.
ACCESSION   AW082879
VERSION     AW082879.1  GI:6038031
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1..451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2581796"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      115 a     91 c     96 g    149 t
ORIGIN

  Query Match             9.6%;  Score 446.2;  DB 10;  Length 451;
  Best Local Similarity   99.3%;  Pred. No. 2.1e-66;
  Matches  448;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4138 CTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACAT 4197
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       451 CTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAAAAACAAAAACAT 392

Qy      4198 CTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAG 4257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       391 CTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAG 332

Qy      4258 ATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCC 4317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       331 ATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCC 272

Qy      4318 AACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAA 4377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       271 AACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAA 212

Qy      4378 GCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAG 4437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       211 GCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAG 152

Qy      4438 GAGACAACAACTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCC 4497
             |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db       151 GAGACAACAACTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTGTGGAAAACTTGTCC 92

Qy      4498 AAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAA 4557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        91 AAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAA 32

Qy      4558 AATTGAATTATGTATATCACTGTTAAAAAAA 4588
             |||||||||||||||||| ||||||||||||
Db        31 AATTGAATTATGTATATCCCTGTTAAAAAAA 1


RESULT 10
BI185454
LOCUS       BI185454                 799 bp    mRNA    linear   EST 10-JUL-2001
DEFINITION  UNL-P-FN-cw-d-01-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
            UNL-P-FN-cw-d-01-0-UNL 3', mRNA sequence.
ACCESSION   BI185454
VERSION     BI185454.1  GI:14659863
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Caetano,A.R., Johnson,R.K. and Pomp,D.
  TITLE     Generation and sequence characterization of a normalized cDNA
            library from swine ovarian follicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Pomp, D
            Department of Animal Science
            University of Nebraska, Lincoln
            Lincoln, NE 68583-0908, USA
            Tel: 402 472 6416
            Fax: 402 472 6362
            Email: dpomp@unl.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message.
            Seq primer: M13 -29
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..799
                     /organism="Sus scrofa"
                     /strain="University of Nebraska, Lincoln Swine Selection
                     Lines"
                     /db_xref="taxon:9823"
                     /clone="UNL-P-FN-cw-d-01-0-UNL"
                     /clone_lib="UNL-P-FN"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
                     library is a normalized library representing porcine
                     ovarian follicles, ranging between 2.0 to 10.0 mm in
                     diameter, collected  during 7 days of the follicular phase
                     of the pig estrous cycle. This  library was derived from
                     the library UNL-P-F2. The tag is a string of  5-6
                     nucleotides present between the Not I site and the
                     oligo-dT track.  The library was constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996.
                     TAG_SEQ=None found"
BASE COUNT      206 a    174 c    241 g    177 t      1 others
ORIGIN

  Query Match             9.0%;  Score 416.8;  DB 13;  Length 799;
  Best Local Similarity   71.1%;  Pred. No. 1.7e-61;
  Matches  564;  Conservative    0;  Mismatches  228;  Indels    1;  Gaps    1;
```

```
Qy   3021 CTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATA 3080
Db      8 CTGCTCAGGGAACCAGAGTCAGACACTATCCCCGTGCAATTCATCATCCTCGGACCCATC 67

Qy   3081 TTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCT 3140
Db     68 AAGCTCTATTATTTCAGAAGAAAATGGTGTTGCCTGCATAGGGAGTGGTCAACTTCGCCT 127

Qy   3141 AGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGG 3200
Db    128 GGTCGATGGAGGTGGTCGTTGTGCTGGGAGAGTAGAGGTCTATCATGAGGGCTCCTGGGG 187

Qy   3201 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGG 3260
Db    188 CACCATCTGTGATGACAGCTGGGACCTGAATGATGCCCATGTGGTGTGCAAACAGCTAAG 247

Qy   3261 CTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCAT 3320
Db    248 CTGTGGATGGGCCATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCAT 307

Qy   3321 CTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCG 3380
Db    308 TTGGCTGGATGAGATAAACTGTAATGGAAAAGAATCTCATATTTGGCAATGCCACTCACA 367

Qy   3381 CGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATT 3440
Db    368 TGGTTGGGGGCGGCACAATTGCAGGCATAAGGAGGATGCAGGAGTCATCTGCTCGGAGTT 427

Qy   3441 CACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT 3500
Db    428 CATGTCTCTGAGACTGATCAGTGAAAACAGCAGAGAGACCTGTGCAGGGCGCCTGGAAGT 487

Qy   3501 CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGG 3560
Db    488 TTTTTACAACGGAGCTTGGGGCAGCGTTGGCAAGAATAGCATGTCTCCAGCCACAGTGGG 547

Qy   3561 CATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATC 3620
Db    548 GGTGGTATGCAGGCAGCTAGGCTGTGCAGACAGAGGGGACATCAGCCCTGCATCTTCAGA 607

Qy   3621 TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTC 3680
Db    608 CAAGACAGTGTCCAGGCACATGTGGGTGGACAATGTTCAGTGTCCTAAAGGACCTGACAC 667

Qy   3681 CATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGAC 3740
Db    668 ACTATGGCAGTGCCCATCATCTCCATGGAAGAAGAGACTGGCCAGCCCCTCAGAGGAGAC 727

Qy   3741 CTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAG 3800
Db    728 ATGGATCACATGTGCCAACAAAATAAGACTTCAAGAAGGA-ANACTAATTGTTCTGGACG 786

Qy   3801 AGTGGAGATCTGG 3813
Db    787 TGTGGAGATCTGG 799
```

```
RESULT 11
AI278132/c
LOCUS       AI278132               429 bp    mRNA    linear   EST 29-JAN-1999
DEFINITION  ql77e02.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1878362
            3', mRNA sequence.
ACCESSION   AI278132
VERSION     AI278132.1  GI:3900400
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1693   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1..429
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1878362"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      108 a     89 c     90 g    142 t
ORIGIN

  Query Match             9.0%;  Score 414.4;  DB 9;  Length 429;
  Best Local Similarity   99.3%;  Pred. No. 5.8e-61;
  Matches  427;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy   4152 GTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGT 4211
Db    428 GTTTATTCTATTTCTCACGTGGTGC--AGTTCAGAAACAAAAACATCTGCCCCTCAGAGT 371

Qy   4212 TTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCT 4271
Db    370 TTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCT 311

Qy   4272 CAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGA 4331
Db    310 CAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGA 251

Qy   4332 AGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACT 4391
Db    250 AGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACT 191

Qy   4392 TTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTT 4451
Db    190 TTAGACTTCCAGGGCTCACCAGCTCAACCTCTAAATATCTTTGAAGGAGACAACAACTTT 131

Qy   4452 TAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTT 4511
Db    130 TAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTT 71

Qy   4512 GAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTA 4571
Db     70 GAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTA 11

Qy   4572 TATCACTGTT 4581
Db     10 TATCACTGTT 1
```

```
RESULT 12
BI183796
LOCUS       BI183796               718 bp    mRNA    linear   EST 10-JUL-2001
DEFINITION  UNL-P-FN-cd-h-08-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
            UNL-P-FN-cd-h-08-0-UNL 3', mRNA sequence.
ACCESSION   BI183796
```

```
VERSION     BI183796.1  GI:14658205
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Caetano,A.R., Johnson,R.K. and Pomp,D.
  TITLE     Generation and sequence characterization of a normalized cDNA
            library from swine ovarian follicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Pomp, D
            Department of Animal Science
            University of Nebraska, Lincoln
            Lincoln, NE 68583-0908, USA
            Tel: 402 472 6416
            Fax: 402 472 6362
            Email: dpomp@unl.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message.
            Seq primer: M13 -29
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..718
                     /organism="Sus scrofa"
                     /strain="University of Nebraska, Lincoln Swine Selection
                     Lines"
                     /db_xref="taxon:9823"
                     /clone="UNL-P-FN-cd-h-08-0-UNL"
                     /clone_lib="UNL-P-FN"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
                     library is a normalized library representing porcine
                     ovarian follicles, ranging between 2.0 to 10.0 mm in
                     diameter, collected  during 7 days of the follicular phase
                     of the pig estrous cycle. This  library was derived from
                     the library UNL-P-F2. The tag is a string of  5-6
                     nucleotides present between the Not I site and the
                     oligo-dT track.  The library was constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996.
                     TAG_SEQ=None found"
BASE COUNT      178 a    155 c    223 g    156 t      6 others
ORIGIN

  Query Match              8.8%;  Score 408.4;  DB 13;  Length 718;
  Best Local Similarity   74.5%;  Pred. No. 4.9e-60;
  Matches  511;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy    3173 TAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCG 3232
           ||||| ||||||| || |||| ||||||||||||||||||||| ||||||||||||   |
Db      12 TAGAGGTCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAATG 71

Qy    3233 ATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTG 3292
           |||||| ||||||||   || ||||| ||||||||    ||||  | |||||||| |  ||||
Db      72 ATGCCCATGTGGTGTGCAAACAGCTGAGCTGTGGATGGGCCATTAATGCCACTGGTTCTG 131

Qy    3293 CTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGG 3352
           |||| |||||||| ||  |||||||||| ||||||||||   | |||||  |  ||||   |
Db     132 CTCATTTTGGGGAAGGAACAGGGCCCATTTGGCTGGATGAGATAAACTGTAATGGAAAAG 191

Qy    3353 AGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGG 3412
           | || ||  | ||||| ||||  || |   || ||||||| |||| | |||||||| ||||
Db     192 AATCTCATATTTGGCAATGCCACTCACATGGTTGGGGGCGGCACAATTGCAGGCATAAGG 251

Qy    3413 AGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAA 3472
           |||| ||||| |||||||||| || ||||    |   | || ||    |||||||      |
Db     252 AGGATGCAGGAGTCATCTGCTCGGAGTTCATGTCTCTCAGACTGATCAGTGAAAACAGCA 311

Qy    3473 CAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCA 3532
            ||||| |||||| ||| |  ||| ||| || || |||||   | ||||||||||| ||||
Db     312 GAGAGACCTGTGCAGGGCGCCTGGNAGTTTTTACAACGGAGCTTGGGGCAGCGTTGGCA 371

Qy    3533 GGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGA 3592
            ||  | |||  |  |||||| ||  ||  | || ||||||||||||||||| || |
Db     372 AGAATAGCATGTCTCCAGCCACAGTGGGGGTGGTATGCAGGCAGCTGGGCTGTGCAGACA 431

Qy    3593 ATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATG 3652
            ||| |   |||||| ||  |||  |   |||||| ||  |   |||||||||||
Db     432 GAGGGGACATCAGCCCTGCATCTTCAGACAAGACAGTGTCCAGGCACATGTGGGTGGACA 491

Qy    3653 ACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGC 3712
           |  |||||||||||||||    | |   | || ||||||||||||  ||   | |||||| ||
Db     492 ATGTTCAGTGTCCTAAAGGACCTGACACCCTATGGCAGTGCCCATCATCTCCATGGAAGA 551

Qy    3713 GAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGC 3772
              ||| |  |||||||   |||| |||||| |||||||||||||   | | ||||||| | |
Db     552 AGAGACTGGCCAGCCCCTCAGAGGAGACATGGATCACATGTGCCAACAAAATAAGACTTC 611

Qy    3773 GTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCA 3832
            | |||  ||||  | || ||| |  | |||||||||||| |||  || ||||||||||
Db     612 AAGAAGGNAACACTAATTGTTCTNGACGTGTGGAGATCTGGTACGGNGGTTCCTGGGGCA 671

Qy    3833 CAGTGTGTGATGACTCCTGGGACCTG 3858
           | |||||||| ||||||||||| ||
Db     672 CTGTGTGTGACGACTCCTGGGACTTG 697


RESULT 13
AI090561/c
LOCUS       AI090561                 411 bp    mRNA    linear   EST 18-AUG-1998
DEFINITION  qa64g07.x1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:1691580 3', mRNA sequence.
ACCESSION   AI090561
VERSION     AI090561.1  GI:3429620
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..411
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1691580"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
BASE COUNT       94 a     81 c     84 g    152 t
ORIGIN
```

```
  Query Match             8.8%;  Score 407.8;  DB 9;  Length 411;
  Best Local Similarity   99.5%;  Pred. No. 7.8e-60;
  Matches  409;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    4192 AAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTC 4251
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 AAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTC 352

Qy    4252 CATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGACAAGAACCTCAGATGAC 4311
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 CATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGACAAGAACCTCAGATGAC 292

Qy    4312 ACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCC 4371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 ACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCC 232

Qy    4372 TCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCT 4431
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db     231 TCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGCTCAACCTCTAAATATCT 172

Qy    4432 TTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAAC 4491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 TTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAAC 112

Qy    4492 TTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGAT 4551
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGAT 52

Qy    4552 AATAAAAATTGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAA 4602
           |||||||||||||||||||||||| |||||||||||||||||||||||||
Db      51 AATAAAAATTGAATTATGTATATCCCTGTTAAAAAAAAAAAAAAAAAAAA 1


RESULT 14
BF132072
LOCUS       BF132072                 483 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  601821079F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4052947 5',
            mRNA sequence.
ACCESSION   BF132072
VERSION     BF132072.1  GI:10971112
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM890  row: n  column: 20
            High quality sequence stop: 483.
FEATURES             Location/Qualifiers
     source          1..483
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4052947"
                     /clone_lib="NIH_MGC_58"
                     /tissue_type="hypernephroma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.35 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      136 a     115 c     106 g     126 t
ORIGIN

  Query Match             8.7%;  Score 401.4;  DB 12;  Length 483;
  Best Local Similarity   93.6%;  Pred. No. 9e-59;
  Matches  452;  Conservative    0;  Mismatches    1;  Indels   30;  Gaps    2;

Qy    4050 TGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGC 4109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGC 60

Qy    4110 ACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCAC 4169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCAC 120

Qy    4170 GTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAG------------------- 4210
           |||||||||||||||||||||||||||||||||||||||||
Db     121 GTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGGTGAGAGCCCCACTAGGAT 180

Qy    4211 ----------TTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGTATAAAATTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 240

Qy    4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGACAAGAACCTCAGATGACACCCCCAAC 4320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAGACCTGCCTCAAGAGAGAGGACCCACATGGACAAGAACCTCAGATGACACCCCCAAC 300

Qy    4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 360

Qy    4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     361 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGCTCAACCTCTAAATATCTTTGAAGGAG 420

Qy    4441 ACAACAACTTTTAAATGAATAAA-GAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAA 4499
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     421 ACAACAACTTTTAAATGAATAAACGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAA 480

Qy    4500 ATA 4502
           |||
Db     481 ATA 483


RESULT 15
BQ024006/c
LOCUS       BQ024006                 606 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-1-BB1p-auu-d-07-0-UI.s1 NCI_CGAP_Pl6 Homo sapiens cDNA clone
            UI-1-BB1p-auu-d-07-0-UI 3', mRNA sequence.
ACCESSION   BQ024006
VERSION     BQ024006.1  GI:19759285
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
```

```
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .606
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BB1p-auu-d-07-0-UI"
                     /clone_lib="NCI_CGAP_Pl6"
                     /tissue_type="Placenta"
                     /dev_stage="Full Term"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Pl6 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_LIB=UI-1-BB1p
                     TAG_TISSUE=placenta human full term
                     TAG_SEQ=AGGAA"
BASE COUNT      148 a    127 c    102 g    229 t
ORIGIN

  Query Match             8.5%;  Score 392.8;  DB 14;  Length 606;
  Best Local Similarity   98.3%;  Pred. No. 2.4e-57;
  Matches  397;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     4196 ATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATG 4255
            | |  ||  ||  |||||||||||||||||||||||||||||||||||||||||||||||
Db      404 AGCATCCATTCTCAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATG 345

Qy     4256 AGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCC 4315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      344 AGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCC 285

Qy     4316 CCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTG 4375
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      284 CCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTG 225

Qy     4376 AAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGA 4435
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      224 AAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGA 165

Qy     4436 AGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGT 4495
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      164 AGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGT 105

Qy     4496 CCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATA 4555
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      104 CCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATA 45

Qy     4556 AAAATTGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAA 4599
            ||||||||||||||||||||||||||||||||||||||||||||
Db       44 AAAATTGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAA 1

Search completed: May 12, 2003, 06:24:01
Job time : 4026 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:          May 12, 2003, 02:56:30 ; Search time 614 Seconds
                                          (without alignments)
                                          9366.838 Million cell updates/sec

Title:           US-09-759-130B-379
Perfect score:   4628
Sequence:        1 gcggccgctcgcgatctaga..........aaacggacgcgtgggtcgac 4628

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters:      1567708

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                 1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                 2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                 3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                 4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                 5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                 6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                 7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                 8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                 9:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                 10:  /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
                 11:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                 12:  /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
                 13:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                 14:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4628	100.0	4628	9	US-10-042-431-9	Sequence 9, Appli
	2	4628	100.0	4628	9	US-09-759-130B-379	Sequence 379, App
	3	4359	94.2	4359	9	US-10-042-431-10	Sequence 10, Appl
	4	4359	94.2	4359	9	US-09-759-130B-380	Sequence 380, App
	5	886.8	19.2	4308	9	US-10-042-431-79	Sequence 79, Appl
	6	886.8	19.2	4308	9	US-09-759-130B-449	Sequence 449, App
c	7	447.4	9.7	449	10	US-09-867-701-10340	Sequence 10340, A
	8	434.6	9.4	473	9	US-09-918-995-32438	Sequence 32438, A
	9	213	4.6	599	9	US-10-040-739-138	Sequence 138, App
	10	199.2	4.3	461	9	US-09-796-692-269	Sequence 269, App
	11	199.2	4.3	461	9	US-09-796-692-7354	Sequence 7354, Ap
	12	199.2	4.3	461	9	US-10-040-862-269	Sequence 269, App
	13	199.2	4.3	461	9	US-10-040-862-7354	Sequence 7354, Ap
	14	197.6	4.3	2562	9	US-09-147-947-5	Sequence 5, Appli
	15	193.2	4.2	2614	9	US-09-147-947-3	Sequence 3, Appli
	16	173.4	3.7	444	10	US-09-925-301-696	Sequence 696, App
	17	169.8	3.7	438	9	US-09-918-995-35432	Sequence 35432, A
	18	149.6	3.2	4344	10	US-09-917-800A-1712	Sequence 1712, Ap
	19	143.2	3.1	697	10	US-09-925-297-237	Sequence 237, App
	20	134.4	2.9	713	9	US-10-040-739-1356	Sequence 1356, Ap
	21	123.4	2.7	1686	9	US-09-905-291A-147	Sequence 147, App
	22	123.4	2.7	1686	9	US-09-902-853-147	Sequence 147, App
	23	123.4	2.7	1686	9	US-09-907-824-147	Sequence 147, App
	24	123.4	2.7	1686	9	US-09-907-841-147	Sequence 147, App
	25	123.4	2.7	1686	9	US-09-904-011-147	Sequence 147, App
	26	123.4	2.7	1686	9	US-09-906-742-147	Sequence 147, App
	27	123.4	2.7	1686	9	US-09-906-838-147	Sequence 147, App
	28	123.4	2.7	1686	9	US-09-907-613-147	Sequence 147, App
	29	123.4	2.7	1686	9	US-09-907-942-147	Sequence 147, App
	30	123.4	2.7	1686	9	US-10-227-884-113	Sequence 113, App
	31	123.4	2.7	1686	9	US-09-904-820-147	Sequence 147, App
	32	123.4	2.7	1686	9	US-09-904-859-147	Sequence 147, App
	33	123.4	2.7	1686	9	US-09-909-204-147	Sequence 147, App
	34	123.4	2.7	1686	9	US-10-230-163-113	Sequence 113, App
	35	123.4	2.7	1686	9	US-09-904-786-147	Sequence 147, App
	36	123.4	2.7	1686	9	US-09-906-646-147	Sequence 147, App
	37	123.4	2.7	1686	9	US-09-906-700-147	Sequence 147, App
	38	123.4	2.7	1686	9	US-09-902-903-147	Sequence 147, App
	39	123.4	2.7	1686	9	US-09-903-749A-147	Sequence 147, App
	40	123.4	2.7	1686	9	US-09-903-786-147	Sequence 147, App
	41	123.4	2.7	1686	9	US-10-218-631-113	Sequence 113, App
	42	123.4	2.7	1686	9	US-10-230-338-113	Sequence 113, App
	43	123.4	2.7	1686	9	US-09-902-736-147	Sequence 147, App
	44	123.4	2.7	1686	9	US-09-904-119-147	Sequence 147, App
	45	123.4	2.7	1686	9	US-09-904-956-147	Sequence 147, App

ALIGNMENTS

```
RESULT 1
US-10-042-431-9
; Sequence 9, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:  SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 4628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-042-431-9

  Query Match             100.0%;  Score 4628;  DB 9;  Length 4628;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4628;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GCGGCCGCTCGCGATCTAGAACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGAT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GCGGCCGCTCGCGATCTAGAACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGAT 60

Qy       61 TTTGGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TTTGGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTC 120

Qy      121 CTGAATTCCTGCTTTCTCATCAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CTGAATTCCTGCTTTCTCATCAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTC 180
```

```
Qy    181 AATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGACAGTGGGGACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGACAGTGGGGACT 240

Qy    241 GTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGT 300

Qy    301 CCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGG 360

Qy    361 CTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAA 420

Qy    421 TGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCC 480

Qy    481 AATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAA 540

Qy    541 TTCCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTG 600

Qy    601 GTGTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTGTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCT 660

Qy    661 GCTGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCTGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720

Qy    721 TGGAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACA 780

Qy    781 TTAACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTAACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGT 840

Qy    841 ATGGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ATGGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGG 900

Qy    901 AACAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AACAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTC 960

Qy    961 GCTGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCTGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020

Qy   1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGT 1080

Qy   1081 CTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTA 1140

Qy   1141 GCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGG 1200

Qy   1201 ACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGA 1260

Qy   1261 TGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATT 1320

Qy   1321 TGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT 1380

Qy   1381 GGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440

Qy   1441 GCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTG 1500

Qy   1501 AAATACCAAGGAGAGTGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AAATACCAAGGAGAGTGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCT 1560

Qy   1561 GTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTT 1620

Qy   1621 AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAAT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAAT 1680

Qy   1681 ATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTG 1740

Qy   1741 ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGC 1800

Qy   1801 TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC 1860

Qy   1861 TGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATT 1920

Qy   1921 GGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980

Qy   1981 TGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGAC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGAC 2040

Qy   2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGG 2100

Qy   2101 CTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTG 2160

Qy   2161 GGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTT 2220

Qy   2221 GAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACAC 2280

Qy   2281 ATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGA 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGA 2340

Qy   2341 TGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2341 TGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCC 2400

Qy   2401 CACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAA 2460

Qy   2461 CATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTG 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 CATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTG 2520

Qy   2521 CTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGA 2580

Qy   2581 AAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTT 2640

Qy   2641 GCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700

Qy   2701 GTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGG 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGG 2760

Qy   2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820

Qy   2821 GAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGA 2880

Qy   2881 GGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGG 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 GGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGG 2940

Qy   2941 AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCAT 3000

Qy   3001 GGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTC 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTC 3060

Qy   3061 GCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTA 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTA 3120

Qy   3121 GAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATC 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATC 3180

Qy   3181 TATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCAC 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCAC 3240

Qy   3241 GTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTT 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTT 3300

Qy   3301 GGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCAC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 GGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCAC 3360

Qy   3361 TTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCA 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCA 3420

Qy   3421 GGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGC 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGC 3480

Qy   3481 TGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAAC 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 TGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAAC 3540

Qy   3541 ATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTT 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 ATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTT 3600

Qy   3601 GTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAG 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAG 3660

Qy   3661 TGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGAGCGAAGAATC 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 TGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGAGCGAAGAATC 3720

Qy   3721 TCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGA 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 TCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGA 3780

Qy   3781 GACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGT 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGT 3840

Qy   3841 GATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCT 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCT 3900

Qy   3901 GCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGAT 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 GCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGAT 3960

Qy   3961 GACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGA 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGA 4020

Qy   4021 CAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAA 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAA 4080

Qy   4081 TCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTT 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 TCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTT 4140

Qy   4141 CTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTG 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 CTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTG 4200

Qy   4201 CCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260

Qy   4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAAC 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAAC 4320

Qy   4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380

Qy   4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440

Qy   4441 ACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAA 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAA 4500

Qy   4501 TAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAAT 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAAT 4560
```

```
Qy   4561 TGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGACGCGT 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 TGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGACGCGT 4620

Qy   4621 GGGTCGAC 4628
          ||||||||
Db   4621 GGGTCGAC 4628


RESULT 2
US-09-759-130B-379
; Sequence 379, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 379
;   LENGTH: 4628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-130B-379

  Query Match             100.0%;  Score 4628;  DB 9;  Length 4628;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4628;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCGGCCGCTCGCGATCTAGAACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCGGCCGCTCGCGATCTAGAACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGAT 60

Qy     61 TTTGGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTGGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTC 120

Qy    121 CTGAATTCCTGCTTTCTCATCAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGAATTCCTGCTTTCTCATCAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTC 180

Qy    181 AATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACT 240

Qy    241 GTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGT 300

Qy    301 CCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAAATTTGG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAAATTTGG 360

Qy    361 CTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAA 420

Qy    421 TGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCC 480

Qy    481 AATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAA 540

Qy    541 TTCCAAGAAAGGTGGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAAGAAAGGTGGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTG 600

Qy    601 GTGTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTGTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCT 660

Qy    661 GCTGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCTGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720

Qy    721 TGGAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACA 780

Qy    781 TTAACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTAACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGT 840

Qy    841 ATGGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGGACCGTATGCCACCATAAGTGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ATGGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGGACCGTATGCCACCATAAGTGG 900

Qy    901 AACAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AACAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTC 960

Qy    961 GCTGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCTGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020

Qy   1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGT 1080

Qy   1081 CTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTA 1140

Qy   1141 GCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGG 1200

Qy   1201 ACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1201 ACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGA 1260

Qy   1261 TGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATT 1320

Qy   1321 TGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT 1380

Qy   1381 GGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440

Qy   1441 GCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTG 1500

Qy   1501 AAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCT 1560

Qy   1561 GTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTT 1620

Qy   1621 AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAAT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAAT 1680

Qy   1681 ATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTG 1740

Qy   1741 ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGC 1800

Qy   1801 TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC 1860

Qy   1861 TGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATT 1920

Qy   1921 GGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980

Qy   1981 TGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGAC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGAC 2040

Qy   2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGG 2100

Qy   2101 CTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTG 2160

Qy   2161 GGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTT 2220

Qy   2221 GAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACAC 2280

Qy   2281 ATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGA 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGA 2340

Qy   2341 TGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCC 2400

Qy   2401 CACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAA 2460

Qy   2461 CATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTG 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 CATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTG 2520

Qy   2521 CTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGA 2580

Qy   2581 AAAGGGAATGGTCTAACTTGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AAAGGGAATGGTCTAACTTGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTT 2640

Qy   2641 GCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700

Qy   2701 GTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGG 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGG 2760

Qy   2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820

Qy   2821 GAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGA 2880

Qy   2881 GGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGG 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 GGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGG 2940

Qy   2941 AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCAT 3000

Qy   3001 GGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTC 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTC 3060

Qy   3061 GCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTA 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTA 3120

Qy   3121 GAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATC 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATC 3180

Qy   3181 TATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCAC 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCAC 3240

Qy   3241 GTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTT 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTT 3300

Qy   3301 GGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCAC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 GGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCAC 3360

Qy   3361 TTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCA 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCA 3420
```

```
Qy   3421 GGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGC 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGC 3480

Qy   3481 TGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAAC 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 TGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAAC 3540

Qy   3541 ATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTT 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 ATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTT 3600

Qy   3601 GTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAG 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAG 3660

Qy   3661 TGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATC 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 TGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATC 3720

Qy   3721 TCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGA 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 TCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGA 3780

Qy   3781 GACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGT 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGT 3840

Qy   3841 GATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCT 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCT 3900

Qy   3901 GCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGAT 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 GCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGAT 3960

Qy   3961 GACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGA 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGA 4020

Qy   4021 CAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAA 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAA 4080

Qy   4081 TCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTT 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 TCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTT 4140

Qy   4141 CTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTG 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 CTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTG 4200

Qy   4201 CCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260

Qy   4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAAC 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAAC 4320

Qy   4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380

Qy   4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440

Qy   4441 ACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAA 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAA 4500

Qy   4501 TAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAAT 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAAT 4560

Qy   4561 TGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGACGCGT 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 TGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGACGCGT 4620

Qy   4621 GGGTCGAC 4628
          ||||||||
Db   4621 GGGTCGAC 4628


RESULT 3
US-10-042-431-10
; Sequence 10, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:  SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4359
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-042-431-10

  Query Match             94.2%;  Score 4359;  DB 9;  Length 4359;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     28 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 87
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 60

Qy     88 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 120

Qy    148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180

Qy    208 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240

Qy    268 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 300

Qy    328 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 360

Qy    388 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 420

Qy    448 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 480
```

```
Qy   508 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 567
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 540

Qy   568 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 627
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600

Qy   628 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 687
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 660

Qy   688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 747
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 720

Qy   748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 807
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 780

Qy   808 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 867
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 840

Qy   868 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 927
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 900

Qy   928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 987
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 960

Qy   988 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1047
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1020

Qy  1048 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1107
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1080

Qy  1108 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1140

Qy  1168 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1227
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1200

Qy  1228 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1287
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1260

Qy  1288 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1347
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1320

Qy  1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1407
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1380

Qy  1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1467
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1440

Qy  1468 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1527
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1500

Qy  1528 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1587
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1560

Qy  1588 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1647
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620

Qy  1648 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1707
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1680

Qy  1708 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1740

Qy  1768 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1827
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1800

Qy  1828 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860

Qy  1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1947
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1920

Qy  1948 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2007
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 1980

Qy  2008 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2040

Qy  2068 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100

Qy  2128 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2187
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2160

Qy  2188 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2247
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220

Qy  2248 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2307
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2280

Qy  2308 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2367
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2340

Qy  2368 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2427
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2400

Qy  2428 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2487
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2460

Qy  2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2547
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2520

Qy  2548 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2607
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2580

Qy  2608 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2667
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2640

Qy  2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2700

Qy       2728 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2760

Qy       2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2820

Qy       2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2880

Qy       2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940

Qy       2968 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000

Qy       3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3060

Qy       3088 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3120

Qy       3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3180

Qy       3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3240

Qy       3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3300

Qy       3328 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3360

Qy       3388 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3420

Qy       3448 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480

Qy       3508 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3540

Qy       3568 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3600

Qy       3628 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3660

Qy       3688 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3720

Qy       3748 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3780

Qy       3808 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3840

Qy       3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3900

Qy       3928 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3960

Qy       3988 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4020

Qy       4048 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4080

Qy       4108 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4140

Qy       4168 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4200

Qy       4228 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4260

Qy       4288 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320

Qy       4348 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4386
              |||||||||||||||||||||||||||||||||||||||
Db       4321 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4359
```

RESULT 4
US-09-759-130B-380
; Sequence 380, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159

```
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 380
;   LENGTH: 4359
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-130B-380

  Query Match             94.2%;  Score 4359;  DB 9;  Length 4359;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      28 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 60

Qy      88 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 120

Qy     148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180

Qy     208 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240

Qy     268 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 300

Qy     328 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 360

Qy     388 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 420

Qy     448 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 480

Qy     508 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 540

Qy     568 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600

Qy     628 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 660

Qy     688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 720

Qy     748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 780

Qy     808 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 840

Qy     868 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 900

Qy     928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 960

Qy     988 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1020

Qy    1048 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1080

Qy    1108 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1167
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1140

Qy    1168 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1200

Qy    1228 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1260

Qy    1288 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1320

Qy    1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1407
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1380

Qy    1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1467
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1440

Qy    1468 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1527
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1500

Qy    1528 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1587
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1560

Qy    1588 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1647
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620

Qy    1648 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1707
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1680

Qy    1708 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1740

Qy    1768 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1827
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1800

Qy    1828 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860

Qy       1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1920

Qy       1948 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 1980

Qy       2008 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2040

Qy       2068 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100

Qy       2128 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2160

Qy       2188 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220

Qy       2248 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2280

Qy       2308 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2340

Qy       2368 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2400

Qy       2428 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2460

Qy       2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2520

Qy       2548 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2580

Qy       2608 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2640

Qy       2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2700

Qy       2728 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2760

Qy       2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2820

Qy       2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2880

Qy       2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940

Qy       2968 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000

Qy       3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3060

Qy       3088 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3120

Qy       3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3180

Qy       3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3240

Qy       3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3300

Qy       3328 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3360

Qy       3388 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3420

Qy       3448 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480

Qy       3508 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3540

Qy       3568 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3600

Qy       3628 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3660

Qy       3688 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3720

Qy       3748 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3780

Qy       3808 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3840

Qy       3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3900

Qy       3928 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3960

Qy       3988 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4020
```

```
Qy     4048 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4107
Db     4021 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4080

Qy     4108 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4167
Db     4081 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4140

Qy     4168 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4227
Db     4141 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4200

Qy     4228 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4287
Db     4201 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4260

Qy     4288 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4347
Db     4261 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320

Qy     4348 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4386
Db     4321 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4359
```

RESULT 5
US-10-042-431-79
; Sequence 79, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 4308
;   TYPE: DNA
;   ORGANISM: Bos sp.
US-10-042-431-79

Query Match 19.2%; Score 886.8; DB 9; Length 4308;
Best Local Similarity 54.9%; Pred. No. 5.5e-233;
Matches 2061; Conservative 0; Mismatches 1527; Indels 168; Gaps 9;

```
Qy      483 TCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATT 542
Db       75 TCTGGAGCTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCA 134

Qy      543 CCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGT 602
Db      135 CCAAGGAGAATGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGT 194

Qy      603 GTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGC 662
Db      195 GTGCAGACAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTTGGGCC 254

Qy      663 TGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTG 722
Db      255 AGGACTTGGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTCAG 314

Qy      723 GAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATT 782
Db      315 TGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGGGA 374

Qy      783 AACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGTAT 842
Db      375 TGCTGGAGTAGTCTGCTCAGGATTTG---TGCGTCTGGCTGGAGGGGATGGACCCTGCTC 431

Qy      843 GGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGGAA 902
Db      432 AGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGATGGGAACTTCAC 491

Qy      903 CAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTCGC 962
Db      492 ACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCTGTGTCTGTCCT 551

Qy      963 TGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTC 1022
Db      552 GGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAGTTCAGGTGTGA 611

Qy     1023 CGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGTCT 1082
Db      612 GGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGGGGCACGTGTCA 671

Qy     1083 TCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTAGC 1142
Db      672 CCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGA---AGTCCGGCTCATGAC 728

Qy     1143 AGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGAC 1202
Db      729 AAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGACAATGGAGAGC 788

Qy     1203 AATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATG 1262
Db      789 GCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGTCAGCTCGGCTG 848

Qy     1263 TCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTG 1322
Db      849 TGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGTGATCAGATCCT 908

Qy     1323 GATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGG 1382
Db      909 AACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGACTGC 968

Qy     1383 AAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGC 1442
Db      969 CCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAACCA 1028

Qy     1443 AGATCTGG---------------------------------------------------- 1450
Db     1029 GATCCAGGTGCTTCCCCAGTGCAACGACTCCGTGTCTCAACCTACAGGCTCTGCGGCCTC 1088

Qy     1451 -----------------------------------ACCTAAGGCTTGTCGGGGCTCATAG 1475
Db     1089 AGAGGACAGCGCCCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGGCGG 1148

Qy     1476 CCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGA 1535
Db     1149 TCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGA 1208

Qy     1536 CAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTAT 1595
Db     1209 CGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGCCCT 1268

Qy     1596 GCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGT 1655
Db     1269 CAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGACAACTT 1328

Qy     1656 TTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCA 1715
Db     1329 GAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCAGCA 1388

Qy     1716 TAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCT 1775
```

```
Db   1389 CAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCAGAGT------TCCTGGCCCT 1442

Qy   1776 GAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACG 1835
Db   1443 CAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAATGGGAC 1502

Qy   1836 GTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCA 1895
Db   1503 CTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGCAGACA 1562

Qy   1896 GCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGG 1955
Db   1563 GCTTGGCTGTGGGGACAGTGGAACCCTCAACTCTTCTGTTGCTCTTAGAGAAGGTTTTAG 1622

Qy   1956 AAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAG 2015
Db   1623 GCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACCTCTCTCTGGCAGTGTCC 1682

Qy   2016 GAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTC 2075
Db   1683 TTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAAGCCTATATCTGGTG--- 1739

Qy   2076 TGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGT 2135
Db   1740 ---TGCAGACAGCAGACAGATCCGCCTGGTGGATGGAGGTGGTCGCTGCTCTGGGAGAGT 1796

Qy   2136 TGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACAT 2195
Db   1797 GGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGACCGCTGGGACCTGGACGA 1856

Qy   2196 TGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCC 2255
Db   1857 TGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTGGACGCCACTGTCTCTTC 1916

Qy   2256 TCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGA 2315
Db   1917 CTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTGAACTGCAGAGGAGAGGA 1976

Qy   2316 AGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATAT 2375
Db   1977 GTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACACAACTGCAATCATCAAGA 2036

Qy   2376 GGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCC 2435
Db   2037 AGATGCAGGAGTCATCTGCTCAGGATTTGTGC------GTCTGGCTGGAGGAGATGGACC 2090

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   2091 CTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACCCCAGTGTCTGATGGAAA 2150

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
Db   2151 CTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGATGTGGCAAGGCTGTGTC 2210

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCA 2615
Db   2211 TGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTCTGGGCTGAAGAGTTCAG 2270

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   2271 GTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTGCCCTGTCCAGGAGGCAC 2330

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735
Db   2331 ATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTACACAGAAGTCCAGCTTAT 2390

Qy   2736 G---AATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGG 2792
Db   2391 GAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAGATCTCTGGACGATGGAG 2450

Qy   2793 CTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAG 2852
Db   2451 AGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTGTCTGTCGTCAGCTCGG 2510

Qy   2853 CTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGT 2912
Db   2511 CTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTGGAAGGAGGTGATCAGAT 2570

Qy   2913 GTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGAC 2972
Db   2571 CTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGAC 2630

Qy   2973 AGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAG 3032
Db   2631 TGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAA 2690

Qy   3033 CCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGT 3092
Db   2691 CCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAACCTGCAGGCTCTGCGGC 2750

Qy   3093 TCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGA 3152
Db   2751 CTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGG 2810

Qy   3153 CAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGA 3212
Db   2811 CGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGA 2870

Qy   3213 TGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGC 3272
Db   2871 TGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGC 2930

Qy   3273 CTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGA 3332
Db   2931 CCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGCTGGACGA 2990

Qy   3333 CCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCA 3392
Db   2991 CCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCG 3050

Qy   3393 GCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAG 3452
Db   3051 GCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCAGAGTTCCTGGCCCTCAG 3110

Qy   3453 GCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGG 3512
Db   3111 G------ATGGTGAGCGAGGACCAGCAGTGTGCTGGGTGGCTGGAGGTTTTCTACAACGG 3164

Qy   3513 GACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAG 3572
Db   3165 GACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACTGTGTCCGTGATCTGCAG 3224

Qy   3573 GCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTC 3632
Db   3225 ACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTTGGTCTCAGGGAAGGTTC 3284

Qy   3633 TGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTG 3692
Db   3285 TAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGATACCTCTCTCTGGCAGTG 3344

Qy   3693 CCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATG 3752
Db   3345 TCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAGGAAGCCTACATCTCATG 3404

Qy   3753 TGAAGATAGAATAA---------------------------------------------- 3766
Db   3405 TGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGCACAGACAGAGAGAAGCT 3464

Qy   3767 --GAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTC 3824
Db   3465 CCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAGGTGTGGCACAACGGCTC 3524

Qy   3825 CTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCA 3884
Db   3525 CTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCTGAGGTGGTGTGTCAGCA 3584
```

```
Qy   3885 GCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGG 3944
Db   3585 GCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCATTTGGCCCTGGAAATGG 3644

Qy   3945 AACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCA 4004
Db   3645 GAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCCTCCCTGTGGGACTGTGT 3704

Qy   4005 CGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTC 4064
Db   3705 TGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGATGCTGGTGTGAGGTGCTC 3764

Qy   4065 TGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4100
Db   3765 TGGTGTAAGGACAACATTGCCCACGACCACAGCAGG 3800
```

RESULT 6
US-09-759-130B-449
; Sequence 449, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J.
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Bovine
US-09-759-130B-449

```
  Query Match             19.2%;  Score 886.8;  DB 9;  Length 4308;
  Best Local Similarity   54.9%;  Pred. No. 5.5e-233;
  Matches 2061;  Conservative    0;  Mismatches 1527;  Indels  168;  Gaps    9;

Qy    483 TCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATT 542
Db     75 TCTGGAGCTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCA 134

Qy    543 CCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGT 602
Db    135 CCAAGGAGAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGT 194

Qy    603 GTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGC 662
Db    195 GTGCAGACAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTTGGGCC 254

Qy    663 TGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTG 722
Db    255 AGGACTTGGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTCAG 314

Qy    723 GAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATT 782
Db    315 TGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGGGA 374

Qy    783 AACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGTAT 842
Db    375 TGCTGGAGTAGTCTGCTCAGGATTTG---TGCGTCTGGCTGGAGGGGATGGACCCTGCTC 431

Qy    843 GGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGGACCGTATGCCACCATAAGTGGAA 902
Db    432 AGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGATGGGAACTTCAC 491

Qy    903 CAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTCGC 962
Db    492 ACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCTGTGTCTGTCCT 551

Qy    963 TGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTC 1022
Db    552 GGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAGTTCAGGTGTGA 611

Qy   1023 CGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGTCT 1082
Db    612 GGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGGGGCACGTGTCA 671

Qy   1083 TCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTAGC 1142
Db    672 CCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGA---AGTCCGGCTCATGAC 728

Qy   1143 AGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGAC 1202
Db    729 AAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGACAATGGAGAGC 788

Qy   1203 AATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATG 1262
Db    789 GCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGTCAGCTCGGCTG 848

Qy   1263 TCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTG 1322
Db    849 TGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGTGATCAGATCCT 908

Qy   1323 GATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGG 1382
Db    909 AACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGACTGC 968

Qy   1383 AAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGC 1442
Db    969 CCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAACCA 1028

Qy   1443 AGATCTGG---------------------------------------------------- 1450
Db   1029 GATCCAGGTGCTTCCCCAGTGCAACGACTCCGTGTCTCAACCTACAGGCTCTGCGGCCTC 1088

Qy   1451 -----------------------------------ACCTAAGGCTTGTCGGGGCTCATAG 1475
Db   1089 AGAGGACAGCGCCCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGGCGG 1148

Qy   1476 CCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGA 1535
```

```
Db   1149 TCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGA 1208

Qy   1536 CAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTAT 1595
Db   1209 CGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGCCCT 1268

Qy   1596 GCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGT 1655
Db   1269 CAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGACAACTT 1328

Qy   1656 TTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCA 1715
Db   1329 GAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCAGCA 1388

Qy   1716 TAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCT 1775
Db   1389 CAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCAGAGT------TCCTGGCCCT 1442

Qy   1776 GAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACG 1835
Db   1443 CAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAATGGGAC 1502

Qy   1836 GTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCA 1895
Db   1503 CTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGCAGACA 1562

Qy   1896 GCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGG 1955
Db   1563 GCTTGGCTGTGGGGACAGTGGAACCCTCAACTCTTCTGTTGCTCTTAGAGAAGGTTTTAG 1622

Qy   1956 AAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAG 2015
Db   1623 GCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACCTCTCTCTGGCAGTGTCC 1682

Qy   2016 GAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTC 2075
Db   1683 TTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAAGCCTATATCTGGTG--- 1739

Qy   2076 TGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGT 2135
Db   1740 ---TGCAGACAGCAGACAGATCCGCCTGGTGGATGGAGGTGGTCGCTGCTCTGGGAGAGT 1796

Qy   2136 TGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACAT 2195
Db   1797 GGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGACCGCTGGGACCTGGACGA 1856

Qy   2196 TGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCC 2255
Db   1857 TGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTGGACGCCACTGTCTCTTC 1916

Qy   2256 TCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGA 2315
Db   1917 CTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTGAACTGCAGAGGAGAGGA 1976

Qy   2316 AGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATAT 2375
Db   1977 GTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACACAACTGCAATCATCAAGA 2036

Qy   2376 GGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCC 2435
Db   2037 AGATGCAGGAGTCATCTGCTCAGGATTTGTGC------GTCTGGCTGGAGGAGATGGACC 2090

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   2091 CTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACCCCAGTGTCTGATGGAAA 2150

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
Db   2151 CTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGATGTGGCAAGGCTGTGTC 2210

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCA 2615
Db   2211 TGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTCTGGGCTGAAGAGTTCAG 2270

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   2271 GTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTGCCCTGTCCAGGAGGCAC 2330

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735
Db   2331 ATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTACACAGAAGTCCAGCTTAT 2390

Qy   2736 G---AATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGG 2792
Db   2391 GAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAGATCTCTGGACGATGGAG 2450

Qy   2793 CTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAG 2852
Db   2451 AGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTGTCTGTCGTCAGCTCGG 2510

Qy   2853 CTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGT 2912
Db   2511 CTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTGGAAGGAGGTGATCAGAT 2570

Qy   2913 GTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGAC 2972
Db   2571 CTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGAC 2630

Qy   2973 AGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAG 3032
Db   2631 TGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAA 2690

Qy   3033 CCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGT 3092
Db   2691 CCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAACCTGCAGGCTCTGCGGC 2750

Qy   3093 TCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGA 3152
Db   2751 CTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGG 2810

Qy   3153 CAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGA 3212
Db   2811 CGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGA 2870

Qy   3213 TGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGC 3272
Db   2871 TGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGC 2930

Qy   3273 CTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGA 3332
Db   2931 CCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGCTGGACGA 2990

Qy   3333 CCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCA 3392
Db   2991 CCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCG 3050

Qy   3393 GCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAG 3452
Db   3051 GCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCAGAGTTCCTGGCCCTCAG 3110

Qy   3453 GCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGG 3512
Db   3111 G------ATGGTGAGCGAGGACCAGCAGTGTGCTGGGTGGCTGGAGGTTTTCTACAACGG 3164

Qy   3513 GACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAG 3572
Db   3165 GACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACTGTGTCCGTGATCTGCAG 3224

Qy   3573 GCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTC 3632
Db   3225 ACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTTGGTCTCAGGGAAGGTTC 3284

Qy   3633 TGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTG 3692
Db   3285 TAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGATACCTCTCTCTGGCAGTG 3344
```

```
Qy   3693 CCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATG 3752
Db   3345 TCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAGGAAGCCTACATCTCATG 3404

Qy   3753 TGAAGATAGAATAA---------------------------------------------- 3766
Db   3405 TGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGCACAGACAGAGAGAAGCT 3464

Qy   3767 --GAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTC 3824
Db   3465 CCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAGGTGTGGCACAACGGCTC 3524

Qy   3825 CTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCA 3884
Db   3525 CTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCTGAGGTGGTGTGTCAGCA 3584

Qy   3885 GCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGG 3944
Db   3585 GCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCATTTGGCCCTGGAAATGG 3644

Qy   3945 AACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCA 4004
Db   3645 GAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCCTCCCTGTGGGACTGTGT 3704

Qy   4005 CGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTC 4064
Db   3705 TGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGATGCTGGTGTGAGGTGCTC 3764

Qy   4065 TGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4100
Db   3765 TGGTGTAAGGACAACATTGCCCACGACCACAGCAGG 3800
```

RESULT 7
US-09-867-701-10340/c
; Sequence 10340, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29 5/26/00
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10340
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-10340

```
  Query Match             9.7%;  Score 447.4;  DB 10;  Length 449;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-112;
  Matches  448;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   4133 GGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAA 4192
Db    449 GGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAA 390

Qy   4193 AACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCC 4252
Db    389 AACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCC 330

Qy   4253 ATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACA 4312
Db    329 ATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACA 270

Qy   4313 CCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCT 4372
Db    269 CCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCT 210

Qy   4373 CTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 4432
Db    209 CTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 150

Qy   4433 TGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACT 4492
Db    149 TGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTGTGGAAAACT 90

Qy   4493 TGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 4552
Db     89 TGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 30

Qy   4553 ATAAAAATTGAATTATGTATATCACTGTT 4581
Db     29 ATAAAAATTGAATTATGTATATCACTGTT 1
```

RESULT 8
US-09-918-995-32438
; Sequence 32438, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076 -?
; PRIOR FILING DATE: 1999-01-20 ?
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32438
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(473)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32438

```
  Query Match             9.4%;  Score 434.6;  DB 9;  Length 473;
  Best Local Similarity   98.9%;  Pred. No. 3.9e-109;
  Matches  437;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   3596 GAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACA 3655
Db     31 GAGAAGNACGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACA 90

Qy   3656 TTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAA 3715
Db     91 TTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAA 150

Qy   3716 GAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTG 3775
Db    151 GAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTG 210

Qy   3776 GAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAG 3835
Db    211 GAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAG 270

Qy   3836 TGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTG 3895
Db    271 TGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTG 330

Qy   3896 GCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGT 3955
Db    331 GCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGT 390

Qy   3956 TGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCT 4015
```

```
Db     391 TGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCT 450

Qy    4016 GGGGACAGAGTGACTGTGGACA 4037
           ||||||||||||||||||||||
Db     451 GGGGACAGAGTGACTGTGGACA 472


RESULT 9
US-10-040-739-138
; Sequence 138, Application US/10040739
; Patent No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;    INFORMATION FOR SEQ ID NO: 138:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 599 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-10-040-739-138

  Query Match             4.6%;  Score 213;  DB 9;  Length 599;
  Best Local Similarity   68.0%;  Pred. No. 4.6e-48;
  Matches  297;  Conservative    0;  Mismatches  140;  Indels    0;  Gaps    0;

Qy    1770 GGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCA 1829
           ||  |||||||| ||||  ||    |||   ||     || ||| |||| ||| |   | ||
Db     161 GGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAAAGTCCA 220

Qy    1830 AGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTG 1889
             |     |||||| || |||||| || |  ||||||| ||    ||||  |     ||||| | ||
Db     221 GGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGTGATTTG 280

Qy    1890 TAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGG 1949
           || |||||||| || ||| || |||||| | |    ||    ||  ||   |    | ||||
Db     281 TAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAGTGCAGG 340

Qy    1950 ATATGGAAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTC 2009
           | ||||  |||||| | ||| ||||||| |||  |||  |||||||||| ||| |||
Db     341 TTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCTTTGGGA 400

Qy    2010 ATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGAT 2069
             |.||  |     |||  ||||||||  ||  | ||||  |||||  |||||||| ||| ||||
Db     401 TTGCAAACATGATGGATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGAC 460

Qy    2070 CTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGG 2129
             ||||||||| ||| ||| ||||||| || |||     |||||| ||||| | ||||||||
Db     461 ATGTTCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGG 520

Qy    2130 AAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAAT 2189
             | ||||||||||  |    | |||  |         | || |     ||||||   |   ||||||||| |
Db     521 GACAGTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACT 580

Qy    2190 GAACATTGCTGAAGTTG 2206
           |||     |||||| | |
Db     581 GAAAGAAGCTGAACTCG 597


RESULT 10
US-09-796-692-269
; Sequence 269, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 269
;   LENGTH: 461
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-269

  Query Match             4.3%;  Score 199.2;  DB 9;  Length 461;
  Best Local Similarity   66.7%;  Pred. No. 2.4e-44;
  Matches  299;  Conservative    0;  Mismatches  148;  Indels    1;  Gaps    1;

Qy     183 TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGT 242
             |||||    |    || ||||| | | | |||  | |||||||| |||| |||||| || ||
Db       5 TGGAGGGAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGT 64
```

```
Qy    243 GTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCC 302
Db     65 GTGTGATGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGG 124

Qy    303 ATTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCT 362
Db    125 AAGTGCTGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTT 184

Qy    363 TGATGATGTTTCCTGTTA-TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAAT 421
Db    185 TGATGATCTTATATGCAACCGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGAT 244

Qy    422 GGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCA 481
Db    245 GGGGAAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAG 304

Qy    482 ATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAAT 541
Db    305 ATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGAT 364

Qy    542 TCCAAGAAAGGTGGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGG 601
Db    365 TCCAAGGAGAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGG 424

Qy    602 TGTGCAGGCAACTAGGATGTCCATCTTC 629
Db    425 CATGCAAGCAACTGGGATGTCCAACTGC 452
```

RESULT 11
US-09-796-692-7354
; Sequence 7354, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7354
;   LENGTH: 461
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-7354

```
  Query Match             4.3%;  Score 199.2;  DB 9;  Length 461;
  Best Local Similarity   66.7%;  Pred. No. 2.4e-44;
  Matches  299;  Conservative    0;  Mismatches  148;  Indels    1;  Gaps    1;

Qy    183 TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGACAGTGGGGACTGT 242
Db      5 TGGAGGGAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGT 64

Qy    243 GTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCC 302
Db     65 GTGTGATGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGG 124

Qy    303 ATTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCT 362
Db    125 AAGTGCTGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTT 184

Qy    363 TGATGATGTTTCCTGTTA-TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAAT 421
Db    185 TGATGATCTTATATGCAACCGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGAT 244

Qy    422 GGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCA 481
Db    245 GGGGAAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAG 304

Qy    482 ATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAAT 541
Db    305 ATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGAT 364

Qy    542 TCCAAGAAAGGTGGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGG 601
Db    365 TCCAAGGAGAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGG 424

Qy    602 TGTGCAGGCAACTAGGATGTCCATCTTC 629
Db    425 CATGCAAGCAACTGGGATGTCCAACTGC 452
```

RESULT 12
US-10-040-862-269
; Sequence 269, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03

```
;  PRIOR APPLICATION NUMBER: US 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: US 09/796,692
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 269
;   LENGTH: 461
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-040-862-269

  Query Match             4.3%;  Score 199.2;  DB 9;  Length 461;
  Best Local Similarity   66.7%;  Pred. No. 2.4e-44;
  Matches  299;  Conservative    0;  Mismatches  148;  Indels    1;  Gaps    1;

Qy    183 TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGACTGT 242
Db      5 TGGAGGGAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGT 64

Qy    243 GTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCC 302
Db     65 GTGTGATGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGG 124

Qy    303 ATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCT 362
Db    125 AAGTGCTGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTT 184

Qy    363 TGATGATGTTTCCTGTTA-TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAAT 421
Db    185 TGATGATCTTATATGCAACCGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGAT 244

Qy    422 GGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCA 481
Db    245 GGGGAAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAG 304

Qy    482 ATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAAT 541
Db    305 ATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGAT 364

Qy    542 TCCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGG 601
Db    365 TCCAAGGAGAATGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGG 424

Qy    602 TGTGCAGGCAACTAGGATGTCCATCTTC 629
Db    425 CATGCAAGCAACTGGGATGTCCAACTGC 452

RESULT 13
US-10-040-862-7354
; Sequence 7354, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/040,862
;  CURRENT FILING DATE:  2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: US 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: US 09/796,692
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7354
;   LENGTH: 461
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-040-862-7354

  Query Match             4.3%;  Score 199.2;  DB 9;  Length 461;
  Best Local Similarity   66.7%;  Pred. No. 2.4e-44;
  Matches  299;  Conservative    0;  Mismatches  148;  Indels    1;  Gaps    1;

Qy    183 TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGACTGT 242
Db      5 TGGAGGGAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGT 64

Qy    243 GTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCC 302
Db     65 GTGTGATGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGG 124

Qy    303 ATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCT 362
Db    125 AAGTGCTGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTT 184

Qy    363 TGATGATGTTTCCTGTTA-TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAAT 421
Db    185 TGATGATCTTATATGCAACCGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGAT 244

Qy    422 GGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCA 481
Db    245 GGGGAAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAG 304

Qy    482 ATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAAT 541
Db    305 ATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGAT 364

Qy    542 TCCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGG 601
Db    365 TCCAAGGAGAATGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGG 424

Qy    602 TGTGCAGGCAACTAGGATGTCCATCTTC 629
Db    425 CATGCAAGCAACTGGGATGTCCAACTGC 452

RESULT 14
US-09-147-947-5
; Sequence 5, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
;  APPLICANT: TSURUOKA, No. US20020160490A1uo
;  APPLICANT: YAMASHIRO, Kyoko
;  APPLICANT: YAMAGUCHI, No. US20020160490A1omi
;  TITLE OF INVENTION: No. US20020160490A1el Serine Protease
```

```
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 2562
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   OTHER INFORMATION:
US-09-147-947-5

  Query Match             4.3%;  Score 197.6;  DB 9;  Length 2562;
  Best Local Similarity   52.2%;  Pred. No. 2.1e-43;
  Matches  505;  Conservative    0;  Mismatches  439;  Indels   24;  Gaps    2;

Qy   1133 TGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAAC 1192
Db    680 TTCGCCTTGCTGGAGGCAGCAGTGTGCATGAAGGCCGGGTGGAGCTCTACCATGCTGGCC 739

Qy   1193 AGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGC 1252
Db    740 AGTGGGGAACCGTTTGTGATGACCAATGGGATGATGCCGATGCAGAAGTGATCTGCAGGC 799

Qy   1253 AGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTA 1312
Db    800 AGCTGGGCCTCAGTGGCATTGCCAAAGCATGGCATCAGGCATATTTTGGGGAAGGGTCTG 859

Qy   1313 GAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCA 1372
Db    860 GCCCAGTTATGTTGGATGAAGTACGCTGCACTGGGAATGAGCTTTCAATTGAGCAGTGTC 919

Qy   1373 CATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTT 1432
Db    920 CAAAGAGCTCCTGGGGAGAGCATAACTGTGGCCATAAAGAAGATGCTGGAGTGTCCTGTA 979

Qy   1433 CTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGAT 1492
Db    980 CCCCTCTAACAGATGGGGTCATCAGACTTGCAGGTGGGAAAGGCAGCCATGAGGGTCGCT 1039

Qy   1493 TGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGACAGATGGAGCACAAGGA 1552
Db   1040 TGGAGGTATATTACAGAGGCCAGTGGGGAACTGTCTGTGATGATGGCTGGACTGAGCTGA 1099

Qy   1553 ATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGA 1612
Db   1100 ATACATACGTGGTTTGTCGACAGTTGGGATTTAAATATGGTAAACAAGCATCTG---CCA 1156

Qy   1613 CCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATG 1672
Db   1157 ACCATTTTGAAGAAAGCACAGGGCCCATATGGTTGGATGACGTCAGCTGCTCAGGAAAGG 1216

Qy   1673 AGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAG 1732
Db   1217 AAACCAGATTTCTTCAGTGTTCCAGGCGACAGTGGGGAAGGCATGACTGCAGCCACCGCG 1276

Qy   1733 AGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTG--------------- 1776
Db   1277 AAGATGTTAGCATTGCCTGCTACCCTGGCGGCGAGGGACACAGGCTCTCTCTGGGTTTTC 1336

Qy   1777 -----AGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAG 1831
Db   1337 CTGTCAGACTGATGGATGGAGAAAATAAGAAAGAAGGACGAGTGGAGGTTTTTATCAATG 1396

Qy   1832 GACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTA 1891
Db   1397 GCCAGTGGGGAACAATCTGTGATGATGGATGGACTGATAAGGATGCAGCTGTGATCTGTC 1456

Qy   1892 GCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGAT 1951
Db   1457 GTCAGCTTGGCTACAAGGGTCCTGCCAGAGCAAGAACCATGGCTTACTTTGGAGAAGGAA 1516

Qy   1952 ATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCAT 2011
Db   1517 AAGGACCCATCCATGTGGATAATGTGAAGTGCACAGGAAATGAGAGGTCCTTGGCTGACT 1576

Qy   2012 GCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCT 2071
Db   1577 GTATCAAGCAAGATATTGGAAGACACAACTGCCGCCACAGTGAAGATGCAGGAGTTATTT 1636

Qy   2072 GTTCTGAT 2079
Db   1637 GTGATTAT 1644


RESULT 15
US-09-147-947-3
; Sequence 3, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
;  APPLICANT: TSURUOKA, No. US20020160490A1uo
;  APPLICANT: YAMASHIRO, Kyoko
;  APPLICANT: YAMAGUCHI, No. US20020160490A1omi
;  TITLE OF INVENTION: No. US20020160490A1el Serine Protease
;  FILE REFERENCE: 001560-349
;  CURRENT APPLICATION NUMBER: US/09/147,947A
;  CURRENT FILING DATE: 1997-03-24
;  EARLIER APPLICATION NUMBER: PCT/JP98/03324
;  EARLIER FILING DATE: 1998-07-24
;  EARLIER APPLICATION NUMBER: JP 9/213969
;  EARLIER FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2614
;   TYPE: DNA
;   ORGANISM: Mouse
;   FEATURE:
;   OTHER INFORMATION:
US-09-147-947-3

  Query Match             4.2%;  Score 193.2;  DB 9;  Length 2614;
  Best Local Similarity   52.0%;  Pred. No. 3.5e-42;
  Matches  500;  Conservative    0;  Mismatches  438;  Indels   24;  Gaps    2;

Qy   1133 TGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAAC 1192
Db    741 TTCGCCTTGTTGGTGGGAACAGTGGGCATGAAGGTCGAGTGGAGCTGTACCACGCTGGCC 800

Qy   1193 AGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGC 1252
Db    801 AGTGGGGGACCATCTGTGACGACCAATGGGACAATGCAGACGCAGACGTCATCTGTAGGC 860

Qy   1253 AGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTA 1312
Db    861 AGCTGGGGCTCAGTGGCATTGCCAAAGCATGGCATCAGGCACATTTTGGGGAAGGATCTG 920

Qy   1313 GAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCA 1372
Db    921 GCCCAATATTGTTGGATGAAGTACGCTGCACCGGAAACGAGCTGTCAATTGAGCAATGTC 980

Qy   1373 CATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTT 1432
Db    981 CAAAGAGTTCCTGGGGCGAACATAACTGTGGCCATAAAGAAGATGCTGGAGTGTCTTGTG 1040

Qy   1433 CTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGAT 1492
Db   1041 TTCCTCTAACAGATGGTGTCATCAGACTGGCAGGAGGAAAAAGTACCCATGAAGGTCGCC 1100

Qy   1493 TGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGACAGATGGAGCACAAGGA 1552
```

```
Db   1101 TGGAGGTCTACTACAAGGGGCAGTGGGGACAGTCTGTGATGATGGCTGGACTGAGATGA 1160

Qy   1553 ATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGA 1612

Db   1161 ACACATACGTGGCTTGTCGACTGCTGGGATTTAAATACGGCAAACA---GTCCTCTGTGA 1217

Qy   1613 CCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATG 1672

Db   1218 ACCATTTTGATGGCAGCAACAGGCCCATATGGCTGGATGACGTCAGCTGCTCAGGAAAAG 1277

Qy   1673 AGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAG 1732

Db   1278 AAGTCAGCTTCATTCAGTGTTCCAGGAGACAGTGGGGAAGGCATGACTGCAGCCATAGAG 1337

Qy   1733 AGGATGTGATTGTAACCTGCTCAGGTGA--------------------TGCAACATGGG 1771

Db   1338 AAGATGTGGGCCTCACCTGCTATCCTGACAGCGATGGACATAGGCTTTCTCCAGGTTTTC 1397

Qy   1772 GCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAG 1831

Db   1398 CCATCAGACTAGTGGATGGAGAGAATAAGAAGGAAGGACGAGTGGAGGTTTTTGTCAATG 1457

Qy   1832 GACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTA 1891

Db   1458 GCCAATGGGGAACAATCTGCGATGACGGATGGACCGATAAGCATGCAGCTGTGATCTGCC 1517

Qy   1892 GCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGAT 1951

Db   1518 GGCAGCTTGGCTATAAGGGTCCTGCCAGAGCAAGGACTATGGCTTATTTGGGGAAGGAA 1577

Qy   1952 ATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCAT 2011

Db   1578 AAGGCCCCATCCACATGGATAATGTGAAGTGCACAGGAAATGAGAAGGCCCTGGCTGACT 1637

Qy   2012 GCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCT 2071

Db   1638 GTGTCAAACAAGACATTGGAAGGCACAACTGCCGCCACAGTGAGGATGCAGGAGTCATCT 1697

Qy   2072 GT 2073

Db   1698 GT 1699
```

Search completed: May 12, 2003, 06:36:37
Job time : 658 secs

```
OM nucleic - nucleic search, using sw model

Run on:         May 12, 2003, 02:54:35 ; Search time 103 Seconds
                                          (without alignments)
                                          13779.612 Million cell updates/sec

Title:          US-09-759-130B-379
Perfect score:  4628
Sequence:       1 gcggccgctcgcgatctaga..........aaacggacgcgtggtcgac 4628

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	197.8	4.3	5802	4	US-09-341-587-4	Sequence 4, Appli
	2	149.6	3.2	4360	1	US-08-470-350B-1	Sequence 1, Appli
	3	140.8	3.0	29598	4	US-09-341-587-6	Sequence 6, Appli
	4	139.2	3.0	28720	4	US-09-341-587-7	Sequence 7, Appli
	5	136	2.9	2001	4	US-09-341-587-2	Sequence 2, Appli
	6	127.6	2.8	2178	3	US-09-034-916-1	Sequence 1, Appli
	7	118.6	2.6	2285	1	US-08-477-674-9	Sequence 9, Appli
	8	118.6	2.6	2285	1	US-08-473-791-9	Sequence 9, Appli
	9	118.6	2.6	2285	2	US-08-316-714-9	Sequence 9, Appli
	10	118.6	2.6	2285	3	US-08-473-673-9	Sequence 9, Appli
	11	108.2	2.3	2262	3	US-09-276-400-3	Sequence 3, Appli
	12	108.2	2.3	2262	4	US-09-448-076-3	Sequence 3, Appli
	13	108.2	2.3	2262	4	US-09-702-572-3	Sequence 3, Appli
	14	108.2	2.3	2920	3	US-09-276-400-1	Sequence 1, Appli
	15	108.2	2.3	2920	4	US-09-448-076-1	Sequence 1, Appli
	16	108.2	2.3	2920	4	US-09-702-572-1	Sequence 1, Appli
c	17	100.8	2.2	410	4	US-09-221-298-83	Sequence 83, Appl
	18	92	2.0	1588	6	5510466-3	Patent No. 5510466
	19	91.6	2.0	2028	1	US-08-453-117-1	Sequence 1, Appli
	20	91.6	2.0	2028	2	US-08-948-222-1	Sequence 1, Appli
	21	91.6	2.0	2028	2	US-08-973-145-1	Sequence 1, Appli
	22	91.6	2.0	2028	5	PCT-US96-08081-1	Sequence 1, Appli
	23	91.6	2.0	2037	1	US-08-154-365-1	Sequence 1, Appli
	24	71	1.5	1560	2	US-08-794-795-5	Sequence 5, Appli
	25	71	1.5	1560	4	US-09-249-200-5	Sequence 5, Appli
	26	69	1.5	1703	2	US-08-794-795-1	Sequence 1, Appli
	27	69	1.5	1703	4	US-09-249-200-1	Sequence 1, Appli
	28	63.4	1.4	1868	1	US-08-392-367B-1	Sequence 1, Appli
	29	63.4	1.4	1868	3	US-08-893-467A-1	Sequence 1, Appli
c	30	57.8	1.2	7218	1	US-08-232-463-14	Sequence 14, Appl
	31	49.8	1.1	1356	4	US-09-276-531-129	Sequence 129, App
	32	49.2	1.1	240	1	US-08-628-417-6	Sequence 6, Appli
	33	48.4	1.0	2447	2	US-09-014-969-14	Sequence 14, Appl
	34	47.8	1.0	1872	4	US-09-801-052-1	Sequence 1, Appli
	35	47.2	1.0	140	1	US-08-628-417-5	Sequence 5, Appli
	36	46.6	1.0	6243	2	US-09-056-075-1	Sequence 1, Appli
	37	46.4	1.0	2110	4	US-09-419-459-1	Sequence 1, Appli
	38	45.4	1.0	1474	4	US-08-821-994-64	Sequence 64, Appl
	39	45.4	1.0	2550	6	5258287-23	Patent No. 5258287
	40	45.2	1.0	2601	4	US-08-569-749-3	Sequence 3, Appli
	41	45.2	1.0	2601	5	PCT-US96-12860-3	Sequence 3, Appli
c	42	44.8	1.0	770	4	US-09-385-982-542	Sequence 542, App
	43	44.8	1.0	1117	4	US-09-247-373B-33	Sequence 33, Appl
	44	44.8	1.0	1602	1	US-08-530-950-3	Sequence 3, Appli
	45	44.8	1.0	1602	3	US-08-888-429A-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-09-341-587-4
; Sequence 4, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
;  APPLICANT: Mollenhauer, Jan
;  TITLE OF INVENTION: Protein Containing an SRCR Domain
;  FILE REFERENCE: 4121-108
;  CURRENT APPLICATION NUMBER: US/09/341,587
;  CURRENT FILING DATE: 1999-08-31
;  EARLIER APPLICATION NUMBER: PCT/DE98/00096
;  EARLIER FILING DATE: 1998-01-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 5802
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-4

  Query Match             4.3%;  Score 197.8;  DB 4;  Length 5802;
  Best Local Similarity   48.1%;  Pred. No. 4e-45;
  Matches 1094;  Conservative    0;  Mismatches  882;  Indels  297;  Gaps    7;

Qy   1446 TCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATA 1505
            | ||| ||| ||||| ||     |     |  |     ||| | ||   || |||||||    |||
Db   1186 TTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATA 1245

Qy   1506 CCAAGGAGAGTGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGT 1565
          || |||      ||||| || |||||| |||| || |||  ||| ||   | ||    ||| ||
Db   1246 CCGAGGCTCCTGGGGCACCGTGTGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGT 1305

Qy   1566 GTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGA 1625
           ||    ||  ||||  |||||    | | | |      |      || |    ||    |||    |
Db   1306 CTGCCGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCGGTTTGGCCA 1365

Qy   1626 AGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTG 1685
            |  |||||||| |||   ||||||||||    |||    ||   ||||||   |   | ||
Db   1366 GGGTTCAGGACCCATTGTCCTGGATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTG 1425

Qy   1686 GGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGT 1745
           |  |||     |||| ||| |||       ||||| |||    ||||| || || |     |||
Db   1426 GAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGT 1485

Qy   1746 AACCTGCTCAG------------------------------------------------- 1756
            | ||||||||
Db   1486 CATCTGCTCAGCTGCCCACTCCTGGTCGACGCCCAGTCCAGACACATTGCCGACCATCAC 1545

Qy   1757 ----------------------GTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAG 1793
                                  |||   |  |    |  ||  ||||||||||||    ||   |
```

```
Db   1546 CTTGCCTGCATCGACAGTAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGG 1605

Qy   1794 CAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGA 1853
Db   1606 TGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGCGA 1665

Qy   1854 TGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTC 1913
Db   1666 TGACAGCTGGGACACCAATGATGCCAATGTCGTCTGCAGGCAACCGGGCTGTGGCTGGGC 1725

Qy   1914 TATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGA 1973
Db   1726 CATGTCAGCCCCAGGAAATGCCCGGTTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGA 1785

Qy   1974 TGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAA 2033
Db   1786 TGTGCGCTGCTCAGGACACGAGTCTTACCCGTGGAGCTGCCCCCACAATGGCTGGCTCTC 1845

Qy   2034 TAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGC----------- 2081
Db   1846 CCACAACTGTGGCCATAGTGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCGGCC 1905

Qy   2082 ------------------------------------------------------ATCGGA 2087
Db   1906 AACACCTAGTCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCTGAATCCAG 1965

Qy   2088 TATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGACGTGAATGT 2147
Db   1966 TTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATA 2025

Qy   2148 CCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGT 2207
Db   2026 CCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGT 2085

Qy   2208 TTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGA 2267
Db   2086 TTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCGGTTTGGCCA 2145

Qy   2268 AAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTG 2327
Db   2146 GGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTATCTGTG 2205

Qy   2328 GGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTT 2387
Db   2206 GAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCTGGTGT 2265

Qy   2388 GATCTGCT--------------CAGCCCACAGGCAGCCC-------------------- 2412
Db   2266 CATCTGCTCAGCTTCCCAGTCCCAGCCGACACCCAGCCCAGACACTTGGCCAACCTCACA 2325

Qy   2413 -------------------------------------AGGCTGGTTGGAGCTGATATGCC 2435
Db   2326 TGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAG 2385

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   2386 GTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACTA 2445

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
Db   2446 CTGGGACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTC 2505

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCA 2615
Db   2506 AGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTGGATGATGTGCG 2565

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   2566 CTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAA 2625

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGAT------------------- 2716
Db   2626 CTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCAGCCGACACC 2685

Qy   2717 -----------------------------------------------------ATACAGA 2723
Db   2686 CAGCCCAGACACTTGGCCAACCTCTCGTGCATCAACAGCAGGATCTGAATCCACTTTGGC 2745

Qy   2724 TGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGG 2783
Db   2746 CCTGAGACTGGTGAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCAAGG 2805

Qy   2784 ACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAG 2843
Db   2806 CTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAACGTGGTCTGCAG 2865

Qy   2844 ACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAG 2903
Db   2866 GCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTC 2925

Qy   2904 TGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTG 2963
Db   2926 AGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTG 2985

Qy   2964 TCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTG 3023
Db   2986 CCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTG 3045

Qy   3024 CACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTT 3083
Db   3046 CTCAGCTGCTCAGTCCCAGTCAACG---CCCAGGCCAGATACTTGGCTGACCACCAACTT 3102

Qy   3084 GTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGT 3143
Db   3103 ACCGGCATTGACAGTAGGATCTGAATCCAGTT------------TGGCTCTGAGGCTGGT 3150

Qy   3144 GGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCAC 3203
Db   3151 GAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTGTATCGAGGCTCCTGGGGAAC 3210

Qy   3204 CATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTG 3263
Db   3211 CGTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTG 3270

Qy   3264 TGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTG 3323
Db   3271 TGGCTGGGCCATGTCGGCCCCAGGAAATGCCCGGTTTGGCCAGGGCTCAGGACCCATTGT 3330

Qy   3324 GCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGG 3383
Db   3331 CCTGGATGATGTGCGCTGCTCAGGGAATGAGTCCTACCTGTGGAGCTGCCCCCACAAAGG 3390

Qy   3384 CTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAG 3436
Db   3391 CTGGCTCACCCACAACTGTGGCCATCACGAAGACGCTGGTGTCATCTGCTCAG 3443
```

RESULT 2
US-08-470-350B-1
; Sequence 1, Application US/08470350B
; Patent No. 5684126
; GENERAL INFORMATION:
; APPLICANT: Li, Xiao
; APPLICANT: Snyder, Solomon H
; TITLE OF INVENTION: Ebnerin: A Secreted von Ebner's Gland
; TITLE OF INVENTION: Protein Associated with Taste Buds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patent In Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,350B
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wolffe, Susan A
;      REGISTRATION NUMBER:  33,568
;      REFERENCE/DOCKET NUMBER:  01107.48790
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4360 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus rattus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  94..3963
US-08-470-350B-1

  Query Match              3.2%;  Score 149.6;  DB 1;  Length 4360;
  Best Local Similarity   53.3%;  Pred. No. 1.2e-31;
  Matches  407;  Conservative    0;  Mismatches  324;  Indels   33;  Gaps    3;

Qy     2722 GATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTT 2781
Db      523 GCTGTGAGGCTGGTGAATGGAGGAGACAGGTGTCGGGGTCGTGTGGAGATCCTTTACCAG 582

Qy     2782 GGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGC 2841
Db      583 GGTTCCTGGGGCACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAACGTGGTGTGC 642

Qy     2842 AGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGA 2901
Db      643 AGGCAGCTGGGCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAGTTTGGACAGGGC 702

Qy     2902 AGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAAC 2961
Db      703 TCTGGGTCCATTGTTCTGGATGACGTGGCCTGTAGAGGACATGAGGCCTATCTGTGGAGC 762

Qy     2962 TGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATC 3021
Db      763 TGCTCCCACCGAGGCTGGCTCTCTCATAACTGTGGACATCAGGAGGATGCTGGAGTCATC 822

Qy     3022 TGCAC-------------AGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGT 3068
Db      823 TGTTCATATTCTCAAACAAGCAGTCCCACACCCGATTCTCAAACAAGCAGTCCCACACCC 882

Qy     3069 ATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCT-TTGATCTGCTTAGAGGACA 3127
Db      883 GGTTGGTGGAACCCCGGGGGCACAAATAACGATGTGTCCTATGGACCCGAACAGACCACA 942

Qy     3128 AACG------------------GCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGG 3168
Db      943 GACGCAACAGATTCTGGTTTGGCTGTGAGGCTGGTGAATGGAGGAGACAGGTGTCAGGGT 1002

Qy     3169 AGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTG 3228
Db     1003 CGTGTGGAGATCCTTTACCAGGGTTCCTGGGGTACCGTGTGTGACGACAGCTGGGACACC 1062

Qy     3229 AGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTC 3288
Db     1063 AAGGATGCCAACGTGGTGTGCAGGCAGCTGGTCTGTGGCTGGGCCTTGTCTGCCCCAGGA 1122

Qy     3289 TCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGA 3348
Db     1123 AGTGCCCACTTTGGACAAGGCTCTGGATCCATTGTTCTGGATGACGTGGCCTGTACAGGA 1182

Qy     3349 ACGGAGTGCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGCAGCACGACTGCAGGCAC 3408
Db     1183 CATGAGGCCTATCTGTGGAGCTGCTCCCACCGAGGCTGGCTCTCTCATAACTGTGGCCAC 1242

Qy     3409 AAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAG 3452
Db     1243 CATGAGGATGCTGGAGTCATCTGTTCAGATGCCCAAACCCAGAG 1286


RESULT 3
US-09-341-587-6
; Sequence 6, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
;  APPLICANT: Mollenhauer, Jan
;  TITLE OF INVENTION: Protein Containing an SRCR Domain
;  FILE REFERENCE: 4121-108
;  CURRENT APPLICATION NUMBER: US/09/341,587
;  CURRENT FILING DATE: 1999-08-31
;  EARLIER APPLICATION NUMBER: PCT/DE98/00096
;  EARLIER FILING DATE: 1998-01-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 29598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-6

  Query Match              3.0%;  Score 140.8;  DB 4;  Length 29598;
  Best Local Similarity   66.4%;  Pred. No. 1.2e-28;
  Matches  202;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy     3133 CTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGC 3192
Db     7049 CTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCGAGGC 7108

Qy     3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAA 3252
Db     7109 TCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGTCTGCAGG 7168

Qy     3253 AAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCA 3312
Db     7169 CAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCA 7228

Qy     3313 GGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGC 3372
Db     7229 GGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCCTACCTGTGGAGCTGC 7288

Qy     3373 CCTTCCCGCGGCTGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGC 3432
Db     7289 CCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGTGAAGACGCTGGTGTCATCTGC 7348

Qy     3433 TCAG 3436
Db     7349 TCAG 7352


RESULT 4
US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
;  APPLICANT: Mollenhauer, Jan
;  TITLE OF INVENTION: Protein Containing an SRCR Domain
;  FILE REFERENCE: 4121-108
;  CURRENT APPLICATION NUMBER: US/09/341,587
;  CURRENT FILING DATE: 1999-08-31
```

```
;  EARLIER APPLICATION NUMBER: PCT/DE98/00096
;  EARLIER FILING DATE: 1998-01-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 28720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-7

  Query Match             3.0%;  Score 139.2;  DB 4;  Length 28720;
  Best Local Similarity   66.1%;  Pred. No. 3.4e-28;
  Matches  201;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy    3133 CTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGC 3192
           || | || ||| |||| |    | | |||  ||  |||| ||| ||      || | |||
Db    2944 CTGAGACTGGTGAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCAAGGC 3003

Qy    3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAA 3252
           | |||||||||| | |||||||||  |||||||    |  |||||| ||||||| ||
Db    3004 TCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAACGTGGTCTGCAGG 3063

Qy    3253 AAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCA 3312
             ||||||||||||    |||| |     ||| | |     ||| || |||||  |||| |||
Db    3064 CAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCA 3123

Qy    3313 GGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGC 3372
           || |||||    |||||||| ||  |||| |||||     |||||   || |||||    |||
Db    3124 GGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGC 3183

Qy    3373 CCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGC 3432
           ||    |    ||||||        ||| ||||  | ||  | || || || |||||||||
Db    3184 CCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGC 3243

Qy    3433 TCAG 3436
           ||||
Db    3244 TCAG 3247
```

RESULT 5

```
US-09-341-587-2
; Sequence 2, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
;  APPLICANT: Mollenhauer, Jan
;  TITLE OF INVENTION: Protein Containing an SRCR Domain
;  FILE REFERENCE: 4121-108
;  CURRENT APPLICATION NUMBER: US/09/341,587
;  CURRENT FILING DATE: 1999-08-31
;  EARLIER APPLICATION NUMBER: PCT/DE98/00096
;  EARLIER FILING DATE: 1998-01-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 2001
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-2

  Query Match             2.9%;  Score 136;  DB 4;  Length 2001;
  Best Local Similarity   65.5%;  Pred. No. 5.1e-28;
  Matches  199;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy    3133 CTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGC 3192
           ||  | || ||| |||| |    | | |||    ||  |||| ||| ||     |     |||
Db     184 CTGAGGCTGGTGAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTGTATCGAGGC 243

Qy    3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAA 3252
           | |||||| ||| | ||||||||| ||||||||    |   |||||| | ||||| ||
Db     244 TCCTGGGGAACCGTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTGCAGG 303

Qy    3253 AAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCA 3312
             ||||||||||||    |||| |     ||| | |     ||| |   |||||  |||| |||
Db     304 CAGCTGGGCTGTGGCTGGGCCATGTCGGCCCCAGGAAATGCCCGGTTTGGCCAGGGCTCA 363

Qy    3313 GGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGC 3372
           || |||||    |||||||| ||  |||| |||| |  |||||| || |||||    |||
Db     364 GGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGGAATGAGTCCTACCTGTGGAGCTGC 423

Qy    3373 CCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGC 3432
           ||    |    ||||||        ||| ||||  | ||  | || ||||| || |||||||||
Db     424 CCCCACAAAGGCTGGCTCACCCACAACTGTGGCCATCACGAAGACGCTGGTGTCATCTGC 483

Qy    3433 TCAG 3436
           ||||
Db     484 TCAG 487
```

RESULT 6

```
US-09-034-916-1
; Sequence 1, Application US/09034916
; Patent No. 6046314
;  GENERAL INFORMATION:
;    APPLICANT:  GEBE, JOHN A.
;    APPLICANT:  SIADAK, ANTHONY W.
;    APPLICANT:  ARUFFO, ALEJANDRO A.
;    TITLE OF INVENTION:  SpALPHA: A NOVEL SCAVENGER RECEPTOR
;    TITLE OF INVENTION:  CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
;    TITLE OF INVENTION:  ANTIBODIES THERETO
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBINS & ASSOCIATES
;      STREET:  90 MIDDLEFIELD ROAD, SUITE 200
;      CITY:  MENLO PARK
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,916
;      FILING DATE:  04-MAR-1998
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/039,956
;      FILING DATE:  06-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  5998-0020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 325-7812
;      TELEFAX:  (650) 325-7823
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2178 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  64..1101
US-09-034-916-1

  Query Match             2.8%;  Score 127.6;  DB 3;  Length 2178;
  Best Local Similarity   51.1%;  Pred. No. 1.3e-25;
  Matches  499;  Conservative    0;  Mismatches  439;  Indels   38;  Gaps    7;

Qy    3134 TCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCT 3193
           | || || ||||  ||     |  ||||||||  ||| | || ||| |   | ||   | |||
Db     131 TGCGGCTGGTGGGGGGCCTCCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCC 190
```

```
Qy   3194 TCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAA 3253
          |||||||| | ||||||||||||||||| | |  || |    ||| |||| |
Db    191 AGTGGGGCACCGTGTGTGATGACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGG 250

Qy   3254 AGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAG 3313
          |||||||||||||| |||  |     || |      | |   | || |        |||
Db    251 AGCTGGGCTGTGGAGCTGCCAGCGGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAG 310

Qy   3314 GGCCCATCTGGCTGGATGACCT-GAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGC 3372
                  ||    | | ||    |||   |  ||   ||    |   || ||
Db    311 AAAAAGAGCAAAAGGTCCTCATCCAATCAGTCAGTTGCACAGGAACAGAAGATACATTGG 370

Qy   3373 CCTTCCCGCGGCTGGGGGCAG--CACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCT 3430
          |             || |  | ||   | || ||   ||   | || || || |||| ||||
Db    371 CTCAGTGTGAGCAAGAAGAAGTTTATGATTGTTCACATGATGAAGATGCTGGGG-CATCG 429

Qy   3431 GCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAA----------------- 3472
            | ||||  || ||    |     |||  ||||   |  |
Db    430 TGTGAGAACCCAGAGAGCTCTTTCTCCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGC 489

Qy   3473 -CAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGC 3531
             | | |    ||    ||  |  ||||||            |||    | |||  ||
Db    490 CCTGGGCATTGCAAGGGACGCGTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGC 549

Qy   3532 AGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAG 3591
            ||    |   | |              |||     | |||||| ||||||||| ||||||   |
Db    550 CAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGGAGG 609

Qy   3592 AATGGAGT---TGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTG 3648
            || | |    |    |  | | |  |           |          |||           ||| ||| ||
Db    610 GCTGTACTGACTCAAAAACGCTGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTG 669

Qy   3649 GATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGG 3708
              | ||     ||  |   |    |       || |   | | ||||   ||||  || |||
Db    670 AGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGG 729

Qy   3709 GAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA------TAGA 3762
          | |    |    | | || || ||    ||||| || ||| ||  ||||||||         |
Db    730 GGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC 789

Qy   3763 ATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGC 3822
            | ||| |     |||||||||| |    ||||||||| || ||||| |     ||||     |||
Db    790 TTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGC 849

Qy   3823 TCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAG 3882
                ||||||  | || |||||||||   |||||             ||||     | ||||| ||   ||
Db    850 GTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAG 909

Qy   3883 CAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACG------CTTCGTTTGGCCAG 3936
          || ||||||||||         | ||   || ||               |              |   | |||||
Db    910 CAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCT 969

Qy   3937 GGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGG 3996
          ||     |||    ||||||| |||||  |   | || |||  |||    |  |||       |  ||    |
Db    970 GGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAG 1029

Qy   3997 GACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTG 4056
              | || ||     || |  |   |||||            |||||       ||| |||||||||    |   ||
Db   1030 CAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTC 1089

Qy   4057 AGGTGCTCTGGACAGT 4072
           |  ||||| ||| |||
Db   1090 ATCTGCTCAGGATAGT 1105
```

RESULT 7
US-08-477-674-9
; Sequence 9, Application US/08477674
; Patent No. 5644035
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert F.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Casipit, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,674
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE:
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-477-674-9

```
  Query Match             2.6%;  Score 118.6;  DB 1;  Length 2285;
  Best Local Similarity   62.1%;  Pred. No. 4.4e-23;
  Matches  187;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy   3134 TCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCT 3193
           | || || |  ||||||| |  |        || | || |||||||   ||   |||
Db    250 TGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAGATCTTCTACAGAGGCC 309

Qy   3194 TCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAA 3253
             |||||||  | |||||  ||   ||||||||||  ||||||  ||| || || |
Db    310 AGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGCGTCGTCTGCCGGG 369

Qy   3254 AGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAG 3313
             |||||||  |      |||  | | ||   || |   |||  ||| ||| | || ||||
Db    370 CCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCCTTCGGGCAAGGATCAG 429

Qy   3314 GGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCC 3373
           | ||||||  ||||| ||   |   | ||||| ||||| ||| ||      ||     | |||
Db    430 GCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCCTCACTGGCCGACTGCA 489

Qy   3374 CTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCT 3433
             ||||  ||||||  | ||  |  ||||||||||| ||    ||||| || || |||||
Db    490 AGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGACGCTGGTGTGGTCTGCA 549

Qy   3434 C 3434
          |
Db    550 C 550
```

RESULT 8
US-08-473-791-9

```
; Sequence 9, Application US/08473791
; Patent No. 5736340
;  GENERAL INFORMATION:
;    APPLICANT:  Koths, Kirston E.
;    APPLICANT:  Halenbeck, Robert F.
;    APPLICANT:  Taylor, Eric W.
;    APPLICANT:  Wang, Alice M.
;    APPLICANT:  Casipit, Clayton L.
;    TITLE OF INVENTION:  Secreted Mac-2-Binding Glycoprotein
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cetus Oncology Corporation
;      STREET:  1400 Fifty-Third Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,791
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/316,714
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/961,404
;      FILING DATE:  15-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goldman, Kenneth M.
;      REGISTRATION NUMBER:  34,174
;      REFERENCE/DOCKET NUMBER:  2595.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 420-3152
;      TELEFAX:  (510) 658-5470
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-473-791-9

  Query Match             2.6%;  Score 118.6;  DB 1;  Length 2285;
  Best Local Similarity   62.1%;  Pred. No. 4.4e-23;
  Matches  187;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy   3134 TCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCT 3193
           | || || |  ||||||| |  |  |        ||  | || |||||||  ||   |||
Db    250 TGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAGATCTTCTACAGAGGCC 309

Qy   3194 TCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAA 3253
            ||||||||  | |||||  ||   |||||||||| ||||||  ||| || || |
Db    310 AGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGCGTCGTCTGCCGGG 369

Qy   3254 AGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAG 3313
            |||||||  |      |||  | | ||   || |   |||  ||| ||| | || ||||
Db    370 CCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCCTTCGGGCAAGGATCAG 429

Qy   3314 GGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCC 3373
          | |||||| |||||| || | ||||| ||||| ||| ||     ||      | |||
Db    430 GCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCCTCACTGGCCGACTGCA 489

Qy   3374 CTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCT 3433
            ||||  |||||| | ||  | ||||||||||| ||   ||||| || ||  |||||
Db    490 AGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGACGCTGGTGTGGTCTGCA 549

Qy   3434 C 3434
          |
Db    550 C 550

RESULT 9
US-08-316-714-9
; Sequence 9, Application US/08316714
; Patent No. 5965382
;  GENERAL INFORMATION:
;    APPLICANT:  Koths, Kirston E.
;    APPLICANT:  Halenbeck, Robert F.
;    APPLICANT:  Taylor, Eric W.
;    APPLICANT:  Wang, Alice M.
;    APPLICANT:  Casipit, Clayton L.
;    TITLE OF INVENTION:  Secreted Mac-2-Binding Glycoprotein
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cetus Oncology Corporation
;      STREET:  1400 Fifty-Third Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/316,714
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/961,404
;      FILING DATE:  15-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goldman, Kenneth M.
;      REGISTRATION NUMBER:  34,174
;      REFERENCE/DOCKET NUMBER:  2595.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 420-3152
;      TELEFAX:  (510) 658-5470
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-316-714-9

  Query Match             2.6%;  Score 118.6;  DB 2;  Length 2285;
  Best Local Similarity   62.1%;  Pred. No. 4.4e-23;
  Matches  187;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy   3134 TCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCT 3193
           | || || |  ||||||| |  |  |        ||  | || |||||||  ||   |||
Db    250 TGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAGATCTTCTACAGAGGCC 309

Qy   3194 TCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAA 3253
            ||||||||  | |||||  ||   |||||||||| ||||||  ||| || || |
Db    310 AGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGCGTCGTCTGCCGGG 369

Qy   3254 AGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAG 3313
            |||||||  |      |||  | | ||   || |   |||  ||| ||| | || ||||
Db    370 CCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCCTTCGGGCAAGGATCAG 429

Qy   3314 GGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCC 3373
          | |||||| |||||| || | ||||| ||||| ||| ||     ||      | |||
Db    430 GCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCCTCACTGGCCGACTGCA 489

Qy   3374 CTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCT 3433
```

```
              ||||  ||||||  | ||  | ||||||||||| ||   ||||| || ||  |||||
Db     490 AGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGACGCTGGTGTGGTCTGCA 549

Qy    3434 C 3434
           |
Db     550 C 550
```

RESULT 10
US-08-473-673-9
; Sequence 9, Application US/08473673
; Patent No. 6069127
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert F.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Casipit, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,673
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE: 29-SEP-1994
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-473-673-9

```
  Query Match             2.6%;  Score 118.6;  DB 3;  Length 2285;
  Best Local Similarity   62.1%;  Pred. No. 4.4e-23;
  Matches  187;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy    3134 TCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCT 3193
            | || || |  ||||||| |  |  |         ||  | || |||||||   ||  |||
Db     250 TGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAGATCTTCTACAGAGGCC 309

Qy    3194 TCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAA 3253
              ||||||||  | |||||  ||   ||||||||||  ||||||  ||| || || |
Db     310 AGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGCGTCGTCTGCCGGG 369

Qy    3254 AGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAG 3313
              ||||||  |      |||  | | ||   || |   |||   ||| |||  | || ||||
Db     370 CCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCCTTCGGGCAAGGATCAG 429

Qy    3314 GGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCC 3373
           | ||||||  ||||||  ||   |  | ||||| |||||  ||| ||      ||     | |||
Db     430 GCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCCTCACTGGCCGACTGCA 489

Qy    3374 CTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCT 3433
              ||||  ||||||  | ||  | ||||||||||| ||   ||||| || ||  |||||
Db     490 AGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGACGCTGGTGTGGTCTGCA 549

Qy    3434 C 3434
           |
Db     550 C 550
```

RESULT 11
US-09-276-400-3
; Sequence 3, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-276-400-3

```
  Query Match             2.3%;  Score 108.2;  DB 3;  Length 2262;
  Best Local Similarity   61.6%;  Pred. No. 3.7e-20;
  Matches  173;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy    3157 CGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGAC 3216
           | ||  |  ||  | || |||||  | |  |  ||    ||||||||||||| |||||
Db     157 CCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGATGAT 216

Qy    3217 GGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
           | ||     |||  | ||||||| |  | || |  ||||||||    || ||||
Db     217 GACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA 276

Qy    3277 AATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTG 3336
                  ||  |  |||  | | |||    ||  |||| | |||||||||||  || ||
Db     277 GGCTGGACCCACAGTGCCAAATATGGCCCTGGAACAGGCCGCATCTGGCTGGACAACTTG 336

Qy    3337 AACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCAC 3396
           | |||||  || || |||          ||  | ||  | ||||| ||||||||| |
Db     337 AGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCCCGGGGCTGGGGGAACAGT 396

Qy    3397 GACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGA 3437
           ||||| | |||| | ||||| || |||||||||||  |||
Db     397 GACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGA 437
```

RESULT 12
US-09-448-076-3
; Sequence 3, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: MNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;  LENGTH: 2262
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-448-076-3

  Query Match             2.3%;  Score 108.2;  DB 4;  Length 2262;
  Best Local Similarity  61.6%;  Pred. No. 3.7e-20;
  Matches  173;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy   3157 CGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGAC 3216
           | ||  |  ||  | || |||||  | |  |  ||     ||||||||||||| |||||
Db    157 CCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGATGAT 216

Qy   3217 GGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
           | ||     |||   | ||||||| |  | || |   ||||||||     || ||||
Db    217 GACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA 276

Qy   3277 AATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTG 3336
                 ||    |  |||  | | |||     ||  |||| | |||||||||||  || ||
Db    277 GGCTGGACCCACAGTGCCAAATATGGCCCTGGAACAGGCCGCATCTGGCTGGACAACTTG 336

Qy   3337 AACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCAC 3396
           | |||||  || || |||          ||      | ||   | ||||| ||||||||| |
Db    337 AGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCCCGGGGCTGGGGGAACAGT 396

Qy   3397 GACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGA 3437
           ||||| | |||| | ||||| || |||||||||||   |||
Db    397 GACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGA 437
```

RESULT 13

```
US-09-702-572-3
; Sequence 3, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
;  APPLICANT:  Khodadoust, Mehran
;  TITLE OF INVENTION:  NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE:  MNI-073
;  CURRENT APPLICATION NUMBER: US/09/702,572
;  CURRENT FILING DATE:  2000-10-31
;  PRIOR APPLICATION NUMBER:  09/276,400
;  PRIOR FILING DATE:  1999-03-25
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-702-572-3

  Query Match             2.3%;  Score 108.2;  DB 4;  Length 2262;
  Best Local Similarity  61.6%;  Pred. No. 3.7e-20;
  Matches  173;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy   3157 CGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGAC 3216
           | ||  |  ||  | || |||||  | |  |  ||     ||||||||||||| |||||
Db    157 CCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGATGAT 216

Qy   3217 GGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
           | ||     |||   | ||||||| |  | || |   ||||||||     || ||||
Db    217 GACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA 276

Qy   3277 AATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTG 3336
                 ||    |  |||  | | |||     ||  |||| | |||||||||||  || ||
Db    277 GGCTGGACCCACAGTGCCAAATATGGCCCTGGAACAGGCCGCATCTGGCTGGACAACTTG 336

Qy   3337 AACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCAC 3396
           | |||||  || || |||          ||      | ||   | ||||| ||||||||| |
Db    337 AGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCCCGGGGCTGGGGGAACAGT 396

Qy   3397 GACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGA 3437
           ||||| | |||| | ||||| || |||||||||||   |||
Db    397 GACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGA 437
```

RESULT 14

```
US-09-276-400-1
; Sequence 1, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
;  APPLICANT: Khodadoust, Mehran
;  TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION: USES THEREFOR
;  FILE REFERENCE: MNI-073
;  CURRENT APPLICATION NUMBER: US/09/276,400
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (143)..(2401)
US-09-276-400-1

  Query Match             2.3%;  Score 108.2;  DB 3;  Length 2920;
  Best Local Similarity  61.6%;  Pred. No. 4.3e-20;
  Matches  173;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy   3157 CGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGAC 3216
           | ||  |  ||  | || |||||  | |  |  ||     ||||||||||||| |||||
Db    299 CCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGATGAT 358

Qy   3217 GGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
           | ||     |||   | ||||||| |  | || |   ||||||||     || ||||
Db    359 GACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA 418

Qy   3277 AATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTG 3336
                 ||    |  |||  | | |||     ||  |||| | |||||||||||  || ||
Db    419 GGCTGGACCCACAGTGCCAAATATGGCCCTGGAACAGGCCGCATCTGGCTGGACAACTTG 478

Qy   3337 AACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCAC 3396
           | |||||  || || |||          ||      | ||   | ||||| ||||||||| |
Db    479 AGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCCCGGGGCTGGGGGAACAGT 538

Qy   3397 GACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGA 3437
           ||||| | |||| | ||||| || |||||||||||   |||
Db    539 GACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGA 579
```

RESULT 15

```
US-09-448-076-1
; Sequence 1, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
;  APPLICANT: Khodadoust, Mehran et al.
;  TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
;  FILE REFERENCE: MNI-073CP
;  CURRENT APPLICATION NUMBER: US/09/448,076
;  CURRENT FILING DATE: 1999-11-23
;  EARLIER APPLICATION NUMBER: 60/117,580
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 09/276,400
;  EARLIER FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (143)..(2401)
US-09-448-076-1

  Query Match             2.3%;  Score 108.2;  DB 4;  Length 2920;
  Best Local Similarity   61.6%;  Pred. No. 4.3e-20;
  Matches  173;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy     3157 CGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGAC 3216
             | ||  |  ||  | || |||||  |  |  |  ||     ||||||||||||| |||||
Db      299 CCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGATGAT 358

Qy     3217 GGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
             | ||     |||   | ||||||| |  | || |    |||||||||     || ||||
Db      359 GACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA 418

Qy     3277 AATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTG 3336
                   ||   |  |||  | | |||    ||  |||| | ||||||||||||  || ||
Db      419 GGCTGGACCCACAGTGCCAAATATGGCCCTGGAACAGGCCGCATCTGGCTGGACAACTTG 478

Qy     3337 AACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCAC 3396
             | |||||  || || |||         ||     | ||  | ||||| ||||||||| |
Db      479 AGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCCCGGGGCTGGGGGAACAGT 538

Qy     3397 GACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGA 3437
             ||||| | |||| | ||||| || ||||||||||||   |||
Db      539 GACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGA 579

Search completed: May 12, 2003, 06:29:49
Job time : 374 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         May 12, 2003, 02:05:59 ; Search time 608 Seconds
                                         (without alignments)
                                         17141.856 Million cell updates/sec

Title:          US-09-759-130B-379
Perfect score:  4628
Sequence:       1 gcggccgctcgcgatctaga..........aaacggacgcgtgggtcgac 4628

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4628	100.0	4628	22	AAF45123	Human TANGO 234 cD
2	4359	94.2	4359	22	AAF45124	Human TANGO 234 OR
3	3459.8	74.8	3670	22	AAS01217	DNA encoding human
4	2761.2	59.7	2800	22	AAS44609	Human full-length
5	1448.2	31.3	3716	22	AAI58649	Human polynucleoti
6	1448.2	31.3	3834	22	AAH98740	Human EST-derived
c 7	1448.2	31.3	3834	22	AAI60435	Human polynucleoti
c 8	1448.2	31.3	3834	22	AAI60436	Human polynucleoti
9	1409.4	30.5	3707	24	AAS94922	Human DNA sequence
10	1339.2	28.9	3811	22	AAI58650	Human polynucleoti
11	886.8	19.2	4308	22	AAF45141	Bovine WC1 ORF. B
12	740	16.0	820	21	AAA43017	Human secreted exp
13	642	13.9	690	22	AAH99258	Human protein enco
14	563	12.2	608	20	AAV89272	EST clone CF118.
15	551.4	11.9	562	22	AAH99326	Human protein enco
c 16	447.4	9.7	449	24	ABL87362	Human ovarian canc
17	404	8.7	450	22	ABA08369	Human M160 precurs
18	379.2	8.2	527	22	AAS44781	Human contig polyn
19	335.4	7.2	415	22	ABA09428	Human Cys-rich sca
20	325.8	7.0	339	24	ABN24521	Human ORFX polynuc
21	319.2	6.9	2697	24	AAS17590	DNA encoding novel
22	213.6	4.6	599	20	AAV90288	EST clone DK64. H
23	213	4.6	599	20	AAV87660	EST clone DK399.
24	207.8	4.5	702	21	AAA43661	Mouse secreted exp
25	199.2	4.3	461	22	AAK54544	Human haematologic
26	197.8	4.3	5802	19	AAV49652	Human SRCR protein
27	197.6	4.3	2562	20	AAX19024	Human serine prote
28	197.6	4.3	3350	20	AAV72589	Human neurotrypsin
29	193.2	4.2	2614	20	AAX19027	Mouse serine prote
30	191.6	4.1	2376	20	AAV72590	Mouse neurotrypsin
31	173.4	3.7	444	21	AAC78302	Human cancer assoc
32	167	3.6	474	22	AAH32999	Human colon cancer
33	162	3.5	1269	22	AAD16353	Human sbg14862SPER
34	149.6	3.2	4344	24	ABK63805	Rat sequence diffe
35	149.6	3.2	4360	18	AAT44068	Rat von Ebner's gl
c 36	143.8	3.1	1165	22	ABA08386	Human M130 antigen
37	143.6	3.1	1589	22	AAH33692	Human colon cancer
38	143.2	3.1	697	21	AAC99009	Human pancreatic c
39	140.8	3.0	29598	19	AAV49654	Human SC2 DNA. Ho
40	140.2	3.0	1659	22	AAD16354	Human sbg14862SPER
41	139.2	3.0	28720	19	AAV49655	Human SC3 DNA. Ho
42	136	2.9	2000	19	AAV49651	Human SRCR protein
43	134.4	2.9	713	20	AAV88878	EST clone HY435.
44	127.8	2.8	1591	19	AAV49556	Human liver cell c
45	127.2	2.7	584	22	AAH98116	Murine 7-transmemb

ALIGNMENTS

```
RESULT 1
AAF45123
ID   AAF45123 standard; cDNA; 4628 BP.
XX
AC   AAF45123;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human TANGO 234 cDNA.
XX
KW   Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS   Homo sapiens.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;    99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
```

```
DR   P-PSDB; AAB66037.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 1; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence 4628 BP; 1174 A; 925 C; 1352 G; 1177 T; 0 other;

  Query Match             100.0%;  Score 4628;  DB 22;  Length 4628;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4628;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGGCCGCTCGCGATCTAGAACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGCCGCTCGCGATCTAGAACTAGTAATGATGCTGCCTCAAAACTCGTGGCATATTGAT 60

Qy         61 TTTGGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTGGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTC 120

Qy        121 CTGAATTCCTGCTTTCTCATCAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGAATTCCTGCTTTCTCATCAGCAGTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTC 180

Qy        181 AATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACT 240

Qy        241 GTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGTGTGATGATGGGTGGAACACTACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGT 300

Qy        301 CCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATTTTCTTTCGCCATGTTTCGTTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGG 360

Qy        361 CTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGATGATGTTTCCTGTTATGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAA 420

Qy        421 TGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCC 480

Qy        481 AATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAA 540

Qy        541 TTCCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTG 600

Qy        601 GTGTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCT 660

Qy        661 GCTGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720

Qy        721 TGGAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACA 780

Qy        781 TTAACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTAACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGT 840

Qy        841 ATGGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATGGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGG 900

Qy        901 AACAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AACAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTC 960

Qy        961 GCTGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020

Qy       1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGT 1080

Qy       1081 CTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTA 1140

Qy       1141 GCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGG 1200

Qy       1201 ACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGA 1260

Qy       1261 TGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATT 1320

Qy       1321 TGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT 1380

Qy       1381 GGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GGAAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440

Qy       1441 GCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTG 1500

Qy       1501 AAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCT 1560

Qy       1561 GTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTT 1620

Qy       1621 AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAAT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAAT 1680

Qy       1681 ATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTG 1740

Qy       1741 ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGTGGCGGCAGCAACCGC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1741 ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGC 1800

Qy    1801 TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC 1860

Qy    1861 TGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 TGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATT 1920

Qy    1921 GGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 GGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980

Qy    1981 TGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGAC 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 TGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGAC 2040

Qy    2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGG 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGG 2100

Qy    2101 CTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTG 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 CTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTG 2160

Qy    2161 GGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTT 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 GGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTT 2220

Qy    2221 GAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACAC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 GAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACAC 2280

Qy    2281 ATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGA 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 ATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGA 2340

Qy    2341 TGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCC 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 TGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCC 2400

Qy    2401 CACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAA 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 CACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAA 2460

Qy    2461 CATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTG 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 CATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTG 2520

Qy    2521 CTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGA 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 CTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGA 2580

Qy    2581 AAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTT 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 AAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTT 2640

Qy    2641 GCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 GCATTATGCCCCATTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700

Qy    2701 GTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGG 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGG 2760

Qy    2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820

Qy    2821 GAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGA 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 GAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGA 2880

Qy    2881 GGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGG 2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 GGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGG 2940

Qy    2941 AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCAT 3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCAT 3000

Qy    3001 GGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTC 3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 GGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTC 3060

Qy    3061 GCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTA 3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTA 3120

Qy    3121 GAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATC 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 GAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATC 3180

Qy    3181 TATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCAC 3240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3181 TATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCAC 3240

Qy    3241 GTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTT 3300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 GTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTT 3300

Qy    3301 GGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCAC 3360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3301 GGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCAC 3360

Qy    3361 TTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCA 3420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3361 TTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCA 3420

Qy    3421 GGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGC 3480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3421 GGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGC 3480

Qy    3481 TGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAAC 3540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3481 TGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAAC 3540

Qy    3541 ATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTT 3600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 ATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTT 3600

Qy    3601 GTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAG 3660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3601 GTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAG 3660

Qy    3661 TGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATC 3720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3661 TGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATC 3720

Qy    3721 TCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGA 3780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3721 TCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGA 3780

Qy    3781 GACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGT 3840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3781 GACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGT 3840

Qy    3841 GATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCT 3900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3841 GATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCT 3900

Qy    3901 GCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGAT 3960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3901 GCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGAT 3960
```

```
Qy   3961 GACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGA 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGA 4020

Qy   4021 CAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAA 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAA 4080

Qy   4081 TCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTT 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 TCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTT 4140

Qy   4141 CTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTG 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 CTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTG 4200

Qy   4201 CCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATG 4260

Qy   4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAAC 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 GAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAAC 4320

Qy   4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCC 4380

Qy   4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440

Qy   4441 ACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAA 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ACAACAACTTTTAAATGAATAAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAA 4500

Qy   4501 TAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAAT 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TAACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAAT 4560

Qy   4561 TGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGACGCGT 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 TGAATTATGTATATCACTGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGACGCGT 4620

Qy   4621 GGGTCGAC 4628
          ||||||||
Db   4621 GGGTCGAC 4628


RESULT 2
AAF45124
ID   AAF45124 standard; cDNA; 4359 BP.
XX
AC   AAF45124;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human TANGO 234 ORF.
XX
KW   Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS   Homo sapiens.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
DR   P-PSDB; AAB66037.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 1; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence 4359 BP; 1060 A; 881 C; 1303 G; 1115 T; 0 other;

  Query Match            94.2%;  Score 4359;  DB 22;  Length 4359;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 4359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     28 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 87
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 60

Qy     88 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 120

Qy    148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180

Qy    208 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240

Qy    268 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 300

Qy    328 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 360

Qy    388 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 420

Qy    448 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 480

Qy    508 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGACTATATGT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGACTATATGT 540

Qy    568 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 627
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600

Qy      628 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 687
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 660

Qy      688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 747
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 720

Qy      748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 807
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 780

Qy      808 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 867
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 840

Qy      868 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 927
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 900

Qy      928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 987
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 960

Qy      988 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1047
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1020

Qy     1048 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1107
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1080

Qy     1108 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1167
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1140

Qy     1168 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1227
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1200

Qy     1228 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1287
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1260

Qy     1288 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1347
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1320

Qy     1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1407
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1380

Qy     1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1467
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1440

Qy     1468 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1527
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1500

Qy     1528 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1587
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1560

Qy     1588 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1647
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620

Qy     1648 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1707
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1680

Qy     1708 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1740

Qy     1768 TGGGGCCTGAGGCTGGTGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1827
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 TGGGGCCTGAGGCTGGTGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1800

Qy     1828 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860

Qy     1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1947
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1920

Qy     1948 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2007
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 1980

Qy     2008 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2040

Qy     2068 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100

Qy     2128 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2187
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2160

Qy     2188 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2247
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220

Qy     2248 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2307
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2280

Qy     2308 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2367
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2340

Qy     2368 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2427
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2400

Qy     2428 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2487
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2460

Qy     2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2547
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2520

Qy     2548 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGCCGAA 2607
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGCCGAA 2580

Qy     2608 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2667
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2640

Qy     2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2700

Qy     2728 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2760
```

```
Qy   2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2820

Qy   2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2880

Qy   2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940

Qy   2968 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000

Qy   3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3060

Qy   3088 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3120

Qy   3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3180

Qy   3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3240

Qy   3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3300

Qy   3328 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3360

Qy   3388 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3420

Qy   3448 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480

Qy   3508 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3540

Qy   3568 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3600

Qy   3628 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3660

Qy   3688 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3720

Qy   3748 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3780

Qy   3808 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3840

Qy   3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3900

Qy   3928 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3960

Qy   3988 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4020

Qy   4048 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4080

Qy   4108 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4140

Qy   4168 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4200

Qy   4228 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4260

Qy   4288 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320

Qy   4348 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4386
          |||||||||||||||||||||||||||||||||||||||
Db   4321 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4359
```

```
RESULT 3
AAS01217
ID   AAS01217 standard; cDNA; 3670 BP.
XX
AC   AAS01217;
XX
DT   04-JUL-2001  (first entry)
XX
DE   DNA encoding human secreted protein, POLY8.
XX
KW   Human secreted protein; therapeutic; diagnostic; human; cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..3475
FT                   /*tag=  a
FT                   /product= "Human secreted protein POLY8"
XX
PN   WO200119856-A2.
XX
PD   22-MAR-2001.
XX
PF   13-SEP-2000; 2000WO-US25106.
XX
PR   13-SEP-1999;   99US-0153629.
PR   16-SEP-1999;   99US-0154520.
PR   20-SEP-1999;   99US-0154762.
PR   13-OCT-1999;   99US-0159231.
PR   12-SEP-2000; 2000US-0659634.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Fernandes E,  Herrmann JL,  Liu X,  Yang M,  Boldog FL;
XX
DR   WPI; 2001-244781/25.
```

```
DR   P-PSDB; AAU00396.
XX
PT   New POLYX polypeptide useful for treating or preventing a POLYX
PT   associated disorder, e.g. cancer -
XX
PS   Claim 9; Page 25-29; 152pp; English.
XX
CC   The sequence represents the coding sequence of human secreted protein,
CC   POLY8. POLYX nucleic acids, polypeptides and antibodies to POLYX can be
CC   used for treating or preventing a POLYX associated disorder in a subject,
CC   preferably a human. These can be used in the manufacture of a medicament
CC   for treating a syndrome associated with a human disease selected from a
CC   POLYX-associated disorder, where the therapeutic is a POLYX polypeptide,
CC   a POLYX nucleotide or a POLYX antibody. They may also be used to screen
CC   for a modulator of activity, or latency, or predisposition to a POLYX
CC   associated disorder, e.g. cancer.
XX
SQ   Sequence 3670 BP; 947 A; 752 C; 1073 G; 898 T; 0 other;

  Query Match             74.8%;  Score 3459.8;  DB 22;  Length 3670;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3464;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1114 GATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTA 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        200 GATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGAAGAGTA 259

Qy       1174 GAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAA 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAA 319

Qy       1234 GCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCT 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCT 379

Qy       1294 AAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAG 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 AAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAG 439

Qy       1354 TCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCA 1413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 TCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGAAGATCA 499

Qy       1414 GATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCAT 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCAT 559

Qy       1474 AGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCAT 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 AGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCAT 619

Qy       1534 GACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCT 1593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 GACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCT 679

Qy       1594 ATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGAC 1653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 ATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGAC 739

Qy       1654 GTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAG 1713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAG 799

Qy       1714 CATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGC 1773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 CATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGC 859

Qy       1774 CTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGA 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 CTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGA 919

Qy       1834 CGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGC 1893
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        920 CGGTGGGGCACAGTGTGTGATGACGGCTGGAACGGTAAAGCTGCAGCTGTGGTGTGTAGC 979

Qy       1894 CAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATAT 1953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 CAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATAT 1039

Qy       1954 GGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGC 2013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 GGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGC 1099

Qy       2014 AGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGT 2073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 AGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGT 1159

Qy       2074 TCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAA 2133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 TCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAA 1219

Qy       2134 GTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAAC 2193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 GTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAAC 1279

Qy       2194 ATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAG 2253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 ATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAG 1339

Qy       2254 CCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGG 2313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1340 CCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGG 1399

Qy       2314 GAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAAT 2373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1400 GAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAAT 1459

Qy       2374 ATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATG 2433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1460 ATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATG 1519

Qy       2434 CCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCT 2493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1520 CCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCT 1579

Qy       2494 GATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATA 2553
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db       1580 GATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTGTGGAGATGCCATA 1639

Qy       2554 TCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTC 2613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1640 TCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTC 1699

Qy       2614 CAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGAC 2673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1700 CAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGAC 1759

Qy       2674 ACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTT 2733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1760 ACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTT 1819

Qy       2734 GTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGC 2793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1820 GTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGC 1879

Qy       2794 TCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGC 2853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1880 TCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGC 1939

Qy       2854 TGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTG 2913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1940 TGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTG 1999

Qy       2914 TGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACA 2973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2000 TGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACA 2059
```

```
Qy    2974 GTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGC 3033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2060 GTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGC 2119

Qy    3034 CTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTT 3093
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2120 CTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTT 2179

Qy    3094 CCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGAC 3153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2180 CCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGAC 2239

Qy    3154 AGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGAT 3213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2240 AGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGAT 2299

Qy    3214 GACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCC 3273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2300 GACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCC 2359

Qy    3274 TTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGCCCATCTGGCTGGATGAC 3333
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2360 TTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGCCCATCTGGCTGGATGAC 2419

Qy    3334 CTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAG 3393
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    2420 CTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAG 2479

Qy    3394 CACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGG 3453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2480 CACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGG 2539

Qy    3454 CTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGG 3513
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    2540 CTCTACAGTGAAACTGAAACAGGGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGG 2599

Qy    3514 ACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGG 3573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2600 ACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGG 2659

Qy    3574 CAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCT 3633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2660 CAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCT 2719

Qy    3634 GGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGC 3693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2720 GGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGC 2779

Qy    3694 CTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGT 3753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2780 CTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGT 2839

Qy    3754 GAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGG 3813
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2840 GAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGG 2899

Qy    3814 CACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTG 3873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2900 CACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTG 2959

Qy    3874 GTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGC 3933
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2960 GTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGC 3019

Qy    3934 CAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTA 3993
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3020 CAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTA 3079

Qy    3994 TGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGC 4053
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3080 TGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGC 3139

Qy    4054 GTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTT 4113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3140 GTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCACTT 3199

Qy    4114 ATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGG 4173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3200 ATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGG 3259

Qy    4174 TGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCT 4233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3260 TGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCT 3319

Qy    4234 CTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGG 4293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3320 CTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGG 3379

Qy    4294 ACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTG 4353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3380 ACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTG 3439

Qy    4354 TTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAG 4413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3440 TTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAG 3499

Qy    4414 ATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCA 4473
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3500 CTCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAATAAAGAGGAAGTCA 3559

Qy    4474 AGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAA 4533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3560 AGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAGGAGAACAGCTAAA 3619

Qy    4534 TTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCACTGTTAAA 4584
           |||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    3620 TTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCCCTGTTAAA 3670
```

```
RESULT 4
AAS44609
ID   AAS44609 standard; DNA; 2800 BP.
XX
AC   AAS44609;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human full-length polynucleotide sequence #34.
XX
KW   Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW   mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW   cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW   nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW   angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW   genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW   cytostatic; antirheumatic; antiarthritic; vulnerary; antiinflammatory;
KW   antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW   neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW   immunostimulant; analgesic; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200164834-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04926.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
PR   17-JUN-2000; 2000US-0597707.
PR   14-JUL-2000; 2000US-0616807.
PR   19-SEP-2000; 2000US-0664641.
XX
PA   (HYSE-) HYSEQ INC.
```

```
XX
PI   Tang YT,  Liu C,  Zhou P,  Asundi V,  Zhang J,  Zhao QA,  Ren F;
PI   Xue AJ,  Yang Y,  Wehrman T,  Wang J,  Ma Y,  Wang D,  Chen R,  Xu C;
PI   Drmanac R;
XX
DR   WPI; 2001-589862/66.
DR   P-PSDB; AAU27709.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis, treatment of
PT   cancer, neurological, inflammatory disorders and for use in arrays for
PT   detection -
XX
PS   Claim 1; SEQ ID No 34; 153pp; English.
XX
CC   Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC   contig polynucleotides encoding polypeptides of the invention. The DNA
CC   and protein sequences are useful for the treatment, diagnosis and
CC   prevention of various types of disorder in a mammalian subject such as a
CC   human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC   such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC   as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC   diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC   disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC   chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC   Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC   disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC   bowel disease. The sequences exhibit activity relating to angiogenesis,
CC   cell proliferation, cell differentiation, stem cell growth factor,
CC   activin or inhibin. Therefore, they can be used to manipulate stem cells
CC   in culture to give rise to neuroepithelial cells that can be used to
CC   augment or replace cells damaged by illness, accidental damage or genetic
CC   disorders. The sequences may also be used for regeneration of bone,
CC   cartilage, tendons and ligaments and in tissue repair and burn healing.
CC   Note: Some sequences for this patent did not form part of the printed
CC   specification, but were obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2800 BP; 697 A; 607 C; 813 G; 683 T; 0 other;

  Query Match             59.7%;  Score 2761.2;  DB 22;  Length 2800;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2766;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1821 GTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGC 1880
              ||   ||  ||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         27 GTGACTTTCAGGACGGTGGGGCACAGCGTGTGATGACGGCTGGAACAGTAAAGCTGCAGC 86

Qy       1881 TGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGTCTGGGAAACGC 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 TGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGTCTGGGAAACGC 146

Qy       1941 TTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGA 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGA 206

Qy       2001 TCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGT 2060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGT 266

Qy       2061 TGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAG 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAG 326

Qy       2121 GTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGG 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGG 386

Qy       2181 CTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAG 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAG 446

Qy       2241 GGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGG 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGG 506

Qy       2301 CTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGAGTGGAAACAGACTGC 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGAGTGGAAACAGACTGC 566

Qy       2361 GTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGT 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGT 626

Qy       2421 TGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTC 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTC 686

Qy       2481 TGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTG 2540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        687 TGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTG 746

Qy       2541 TGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTG 2600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTG 806

Qy       2601 GGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCA 2660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCA 866

Qy       2661 ACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATAC 2720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATAC 926

Qy       2721 AGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 2780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 AGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 986

Qy       2781 TGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATG 2840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 TGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATG 1046

Qy       2841 CAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAG 2900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAG 1106

Qy       2901 AAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAA 2960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 AAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAA 1166

Qy       2961 CTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGAT 3020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 CTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGAT 1226

Qy       3021 CTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATA 3080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 CTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATA 1286

Qy       3081 TTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCT 3140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCT 1346

Qy       3141 AGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGG 3200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 AGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGG 1406

Qy       3201 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGG 3260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGG 1466

Qy       3261 CTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCAT 3320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 CTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCAT 1526

Qy       3321 CTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCG 3380
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
```

```
Db   1527 CTGGCTGGATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCG 1586

Qy   3381 CGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATT 3440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1587 CGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATT 1646

Qy   3441 CACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT 3500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 CACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT 1706

Qy   3501 CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGG 3560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGG 1766

Qy   3561 CATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATC 3620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1767 CATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATC 1826

Qy   3621 TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTC 3680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1827 TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTC 1886

Qy   3681 CATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGAC 3740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1887 CATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGAC 1946

Qy   3741 CTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAG 3800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1947 CTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAG 2006

Qy   3801 AGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGC 3860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2007 AGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGC 2066

Qy   3861 CGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGA 3920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2067 CGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGA 2126

Qy   3921 CGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAA 3980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2127 CGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAA 2186

Qy   3981 TGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAA 4040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2187 TGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAA 2246

Qy   4041 GGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2247 GGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 2306

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2307 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 2366

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2367 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 2426

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2427 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 2486

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAG 4340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2487 GGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAG 2546

Qy   4341 CGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTC 4400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2547 CGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAATGACTTTAGACTTC 2606

Qy   4401 CAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAAT 4460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2607 CAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAACAACTTTTAAATGAAT 2666

Qy   4461 AAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAG 4520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2667 AAAGAGGAAGTCAAGTTGCCCTATGGAAAACTTGTCCAAATAACATTTCTTGAACAATAG 2726

Qy   4521 GAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCACTGT 4580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2727 GAGAACAGCTAAATTGATAAAGACTGGTGATAATAAAAATTGAATTATGTATATCACTGT 2786

Qy   4581 TAAAAAAAAAAAAA 4594
          ||||||||||||||
Db   2787 TAAAAAAAAAAAAA 2800


RESULT 5
AAI58649
ID   AAI58649 standard; cDNA; 3716 BP.
XX
AC   AAI58649;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 852.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39493.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 852; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
```

```
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3716 BP; 993 A; 730 C; 1060 G; 933 T; 0 other;

  Query Match            31.3%;  Score 1448.2;  DB 22;  Length 3716;
  Best Local Similarity  65.4%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches 1128;  Indels   24;  Gaps    3;

Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACTTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy   1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
Db    287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy   1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
Db    347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy   1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
Db    407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy   1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
Db    467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy   1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
Db    527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy   1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGACAG 1538
Db    587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db    647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy   1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db    707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1846

Qy   2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db   1847 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1906

Qy   2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db   1907 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966

Qy   2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db   1967 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 2026

Qy   2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db   2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 2086

Qy   2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
```

```
Db   2087 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 2146

Qy   3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
Db   2147 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 2206

Qy   3099 GGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
Db   2207 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 2266

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db   2267 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 2326

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   2327 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 2386

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   2387 TGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 2446

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGA 3398
Db   2447 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAA 2506

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   2507 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 2566

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   2567 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTG 2626

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   2627 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 2686

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   2687 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 2746

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   2747 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 2806

Qy   3699 TGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   2807 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 2866

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db   2867 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 2926

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
Db   2927 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 2986

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
Db   2987 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 3046

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
Db   3047 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 3106

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
Db   3107 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 3166

Qy   4059 GTGCTCTGGACAGTCGCTGAA-----------------ATCACTGAATGCCTCCTCAGG 4100
Db   3167 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 3226

Qy   4101 TCATTTAGCACTTATTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
Db   3227 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 3286

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
Db   3287 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAG 3343

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
Db   3344 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 3403

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGAC 4311
Db   3404 TGATCTGGACCTAATGAATTCCTCAGGAGGC 3434


RESULT 6
AAH98740
ID   AAH98740 standard; cDNA; 3834 BP.
XX
AC   AAH98740;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Human EST-derived coding sequence SEQ ID NO: 597.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition; ss.
XX
OS   Homo sapiens.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   P-PSDB; AAM24081.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 1; Page 584-585; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a cDNA
CC   of the invention.
XX
SQ   Sequence 3834 BP; 1017 A; 756 C; 1082 G; 979 T; 0 other;

  Query Match             31.3%;  Score 1448.2;  DB 22;  Length 3834;
  Best Local Similarity   65.4%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches 1128;  Indels   24;  Gaps    3;
```

```
Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy   1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
Db    287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy   1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
Db    347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy   1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
Db    407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy   1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
Db    467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy   1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
Db    527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy   1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538
Db    587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db    647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy   1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db    707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1846

Qy   2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db   1847 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1906

Qy   2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db   1907 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966

Qy   2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db   1967 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 2026

Qy   2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db   2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 2086

Qy   2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db   2087 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 2146

Qy   3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
Db   2147 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 2206

Qy   3099 GGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
Db   2207 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 2266

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
```

```
Db   2267 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 2326

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   2327 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 2386

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   2387 TGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 2446

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGCAGCACGA 3398
Db   2447 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGCAGCAAAA 2506

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   2507 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 2566

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   2567 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTG 2626

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   2627 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 2686

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   2687 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 2746

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   2747 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 2806

Qy   3699 TGCCCCATGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   2807 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 2866

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db   2867 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 2926

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
Db   2927 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 2986

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
Db   2987 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 3046

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
Db   3047 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 3106

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
Db   3107 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 3166

Qy   4059 GTGCTCTGGACAGTCGCTGAA-----------------ATCACTGAATGCCTCCTCAGG 4100
Db   3167 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 3226

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
Db   3227 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 3286

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
Db   3287 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA--GCGGCTTGCAGTTTCCTCAAG 3343

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280

Db   3344 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 3403

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGAC 4311
Db   3404 TGATCTGGACCTAATGAATTCCTCAGGAGGC 3434
```

RESULT 7
AAI60435/c
ID   AAI60435 standard; cDNA; 3834 BP.
XX
AC   AAI60435;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4424.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41279.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4424; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

```
Query Match            31.3%;  Score 1448.2;  DB 22;  Length 3834;
Best Local Similarity  65.4%;  Pred. No. 0;
Matches 2179;  Conservative    0;  Mismatches 1128;  Indels   24;  Gaps    3;

Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db   3728 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 3669

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db   3668 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 3609

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
Db   3608 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 3549

Qy   1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
Db   3548 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 3489

Qy   1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
Db   3488 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 3429

Qy   1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
Db   3428 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 3369

Qy   1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
Db   3368 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 3309

Qy   1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
Db   3308 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 3249

Qy   1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538
Db   3248 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 3189

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db   3188 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 3129

Qy   1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db   3128 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 3069

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db   3068 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 3009

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db   3008 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 2949

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db   2948 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 2889

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db   2888 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 2829

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   2828 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 2769

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   2768 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 2709

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   2708 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 2649

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   2648 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 2589

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   2588 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 2529

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   2528 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 2469

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   2468 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 2409

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   2408 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 2349

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   2348 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 2289

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   2288 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 2229

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   2228 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 2169

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   2168 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 2109

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   2108 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 2049

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db   2048 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1989

Qy   2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db   1988 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1929

Qy   2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db   1928 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1869

Qy   2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db   1868 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 1809

Qy   2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db   1808 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 1749

Qy   2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db   1748 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 1689

Qy   3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
Db   1688 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 1629

Qy   3099 GGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
```

```
Db   1628 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 1569

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db   1568 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 1509

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   1508 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 1449

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   1448 TGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 1389

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGA 3398
Db   1388 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAA 1329

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   1328 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 1269

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   1268 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTG 1209

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   1208 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 1149

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   1148 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 1089

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   1088 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 1029

Qy   3699 TGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   1028 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 969

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db    968 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 909

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
Db    908 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 849

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
Db    848 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 789

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
Db    788 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 729

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
Db    728 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 669

Qy   4059 GTGCTCTGGACAGTCGCTGAA------------------ATCACTGAATGCCTCCTCAGG 4100
Db    668 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 609

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
Db    608 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 549

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
Db    548 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAG 492

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
Db    491 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 432

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGAC 4311
Db    431 TGATCTGGACCTAATGAATTCCTCAGGAGGC 401
```

RESULT 8
AAI60436/c
ID   AAI60436 standard; cDNA; 3834 BP.
XX
AC   AAI60436;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4425.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41280.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4425; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.

```
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

  Query Match             31.3%;  Score 1448.2;  DB 22;  Length 3834;
  Best Local Similarity   65.4%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches 1128;  Indels   24;  Gaps    3;

Qy      1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db      3728 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 3669

Qy      1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db      3668 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 3609

Qy      1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
Db      3608 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 3549

Qy      1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
Db      3548 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 3489

Qy      1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
Db      3488 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGCTAATTCCAG 3429

Qy      1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
Db      3428 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 3369

Qy      1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
Db      3368 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 3309

Qy      1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
Db      3308 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 3249

Qy      1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538
Db      3248 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 3189

Qy      1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db      3188 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 3129

Qy      1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db      3128 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 3069

Qy      1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db      3068 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 3009

Qy      1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db      3008 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 2949

Qy      1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db      2948 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 2889

Qy      1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db      2888 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 2829

Qy      1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db      2828 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 2769

Qy      1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db      2768 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 2709

Qy      2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db      2708 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 2649

Qy      2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db      2648 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 2589

Qy      2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db      2588 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 2529

Qy      2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db      2528 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 2469

Qy      2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db      2468 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 2409

Qy      2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db      2408 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 2349

Qy      2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db      2348 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 2289

Qy      2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db      2288 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 2229

Qy      2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db      2228 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 2169

Qy      2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db      2168 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 2109

Qy      2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db      2108 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 2049

Qy      2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
Db      2048 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGAA 1989

Qy      2739 TGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACT 2798
Db      1988 TGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCT 1929

Qy      2799 GTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGG 2858
Db      1928 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1869

Qy      2859 GACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGG 2918
Db      1868 AGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAG 1809

Qy      2919 ACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCT 2978
Db      1808 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCT 1749

Qy      2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db      1748 AGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTC 1689

Qy      3039 CCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGA 3098
```

```
Db   1688 CCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGA 1629

Qy   3099 GGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCG 3158
Db   1628 AGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCG 1569

Qy   3159 CTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db   1568 CTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAG 1509

Qy   3219 CTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAA 3278
Db   1508 CTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAA 1449

Qy   3279 TGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db   1448 TGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAA 1389

Qy   3339 CTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGA 3398
Db   1388 ATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAA 1329

Qy   3399 CTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTA 3458
Db   1328 TTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGAC 1269

Qy   3459 CAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db   1268 CAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTG 1209

Qy   3519 GGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCT 3578
Db   1208 GGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCT 1149

Qy   3579 GGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db   1148 GGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCC 1089

Qy   3639 CATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db   1088 CATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATC 1029

Qy   3699 TGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGA 3758
Db   1028 ATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAA 969

Qy   3759 TAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGC 3818
Db    968 CAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGG 909

Qy   3819 AGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878
Db    908 AGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTG 849

Qy   3879 TCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGG 3938
Db    848 TCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGG 789

Qy   3939 AACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGA 3998
Db    788 GACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGA 729

Qy   3999 CTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAG 4058
Db    728 TTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAA 669

Qy   4059 GTGCTCTGGACAGTCGCTGAA-----------------ATCACTGAATGCCTCCTCAGG 4100
Db    668 TTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCG 609

Qy   4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCT 4160
Db    608 TCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGC 549

Qy   4161 ATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAG 4220
Db    548 ATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAG 492

Qy   4221 AAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGA 4280
Db    491 AGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGA 432

Qy   4281 GGACCCACATGGGACAAGAACCTCAGATGAC 4311
Db    431 TGATCTGGACCTAATGAATTCCTCAGGAGGC 401
```

RESULT 9
AAS94922
ID   AAS94922 standard; DNA; 3707 BP.
XX
AC   AAS94922;
XX
DT   14-FEB-2002 (first entry)
XX
DE   Human DNA sequence #177 expressed during foam cell differentiation.
XX
KW   Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW   cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177389-A2.
XX
PD   18-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US11128.
XX
PR   05-APR-2000; 2000US-195106P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI   Tai J;
XX
DR   WPI; 2002-010925/01.
XX
PT   Composition useful for diagnosis of conditions, disorders or diseases
PT   associated with atherosclerosis, comprises several polynucleotides that
PT   are differentially expressed in foam cell development -
XX
PS   Claim 1; Page 232-233; 315pp; English.
XX
CC   The present invention relates to the isolation of human polynucleotide
CC   sequences that are differentially expressed during foam cell
CC   differentiation. The polynucleotide sequences of the invention or a
CC   composition comprising these polynucleotides are useful as a high
CC   throughput method for detecting altered expression of one or more
CC   polynucleotides in a sample. The polynucleotides can be used in the
CC   diagnosis of disorders associated with foam cell development such as
CC   atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC   coronary artery disease. The polynucleotide sequences can also be used
CC   as PCR primers and probes. The polynucleotides of the invention are also
CC   useful in gene therapy. AAS94746-AAS95021 represent the human
CC   polynucleotide sequences of the invention which are differentially
CC   expressed during foam cell differentiation.
XX
SQ   Sequence 3707 BP; 984 A; 731 C; 1062 G; 930 T; 0 other;

Query Match 30.5%; Score 1409.4; DB 24; Length 3707;
Best Local Similarity 65.2%; Pred. No. 0;
Matches 2176; Conservative 0; Mismatches 1131; Indels 28; Gaps 6;

```
Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
```

```
Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGT-TCAGGGAGAGTAGAGGTGA 1180
Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGGAGAGTGGAAGTGA 286

Qy   1181 GAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTG 1240
Db    287 AAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTG 346

Qy   1241 TGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTA 1300
Db    347 TGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCA 406

Qy   1301 GTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTC 1360
Db    407 GTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTC 466

Qy   1361 TCTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATG 1417
Db    467 TTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATG 526

Qy   1418 CTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCC 1477
Db    527 CTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATA 586

Qy   1478 CCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACA 1537
Db    587 TGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATA 646

Qy   1538 GATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGC 1597
Db    647 ACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCA 706

Qy   1598 ATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTT 1657
Db    707 GTTTCTCTGGTTCATCTAATTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTA 766

Qy   1658 CTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATA 1717
Db    767 TATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATA 826

Qy   1718 ATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGA 1777
Db    827 ACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGA 886

Qy   1778 GGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGT 1837
Db    887 GACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAAT 946

Qy   1838 GGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGC 1897
Db    947 GGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAAC 1006

Qy   1898 TGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAA 1957
Db   1007 TGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGAC 1066

Qy   1958 AAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGA 2017
Db   1067 ACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAAC 1126

Qy   2018 ACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTG 2077
Db   1127 ACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTG 1186

Qy   2078 ATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTG 2137
Db   1187 ATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTG 1246

Qy   2138 AGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTG 2197
Db   1247 AGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAG 1306

Qy   2198 CTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTC 2257
Db   1307 CTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGT 1366

Qy   2258 ATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAG 2317
Db   1367 ACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAA 1426

Qy   2318 CCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGG 2377
Db   1427 CTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAG 1486

Qy   2378 AAGCAAGTTTGATCTGCTCAG--CCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCC 2435
Db   1487 AAGCCAAAATTACCTGCTCAGGCCCACAGGGGAACCCAGACTGGTTGGAGGGGACATTCC 1546

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   1547 CTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGCTCCATCTGTGATTCAGA 1606

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
Db   1607 CTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTC 1666

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAGTTCCA 2615
Db   1667 TATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCA 1726

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   1727 GTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAAC 1786

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735
Db   1787 TTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGT 1846

Qy   2736 GAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTC 2795
Db   1847 GAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATC 1906

Qy   2796 ACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTG 2855
Db   1907 CCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATG 1966

Qy   2856 TGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTG 2915
Db   1967 TGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTG 2026

Qy   2916 GGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGACAGT 2975
Db   2027 GAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGC 2086

Qy   2976 TCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCT 3035
Db   2087 TCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCA 2146

Qy   3036 GACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGTTCC 3095
Db   2147 GTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCTACCATTCC 2206

Qy   3096 AGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAG 3155
Db   2207 AGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGG 2266

Qy   3156 CCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGA 3215
Db   2267 TCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGA 2326
```

```
Qy    3216 CGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTT 3275
            | ||||||||||||| ||||||||||||| ||     | ||||||||||||||| |||| |
Db    2327 CAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCAT 2386

Qy    3276 CAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCT 3335
            ||||||| |  |||||||| |||||||| ||  |||||||||||||||||||||   |
Db    2387 TAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGAT 2446

Qy    3336 GAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCA 3395
           ||| ||||  ||||  || |||| | | ||||||||| |||  | ||||||||||||||
Db    2447 GAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCA 2506

Qy    3396 CGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCT 3455
             | |||||||||||||||| || || || |||||||||||||||     |  |||| ||
Db    2507 AAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACT 2566

Qy    3456 CTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGAC 3515
             ||||||  |     | ||||  ||||||  |||  |  ||||||  || ||  ||   |
Db    2567 GACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGC 2626

Qy    3516 CTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCA 3575
            |||||||  || |||| ||| | |||  |     | |||  |  ||   | ||||||||||
Db    2627 TTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCA 2686

Qy    3576 GCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGG 3635
           ||||||||||  || || ||       ||| ||  ||  |||||   ||| |    ||
Db    2687 GCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCAT 2746

Qy    3636 TTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATAT-GGCAGTGCC 3694
            |  |||||||||||  |  ||||||||| |||     | |  | |  |  | |||||||||
Db    2747 TCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGGCAGTGCC 2806

Qy    3695 TGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTG 3754
             ||  | ||||||||     |||  |   ||||||  |  ||  |||||||||||||||||
Db    2807 CATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTG 2866

Qy    3755 AAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGC 3814
            |  |  |  |||||| |  |   | ||||  |||   |  |||||  |  ||||||||||||
Db    2867 ACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGC 2926

Qy    3815 ACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGG 3874
            | | ||| |||||||| ||||||||||||||||| |||||| ||| ||| ||   | ||||
Db    2927 ATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGG 2986

Qy    3875 TGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCC 3934
           |||||| || || |||||||| | ||| ||    ||  | |  || ||   |||||| |
Db    2987 TGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTC 3046

Qy    3935 AGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTAT 3994
           |||| |||||| | || ||| |  ||||  ||    ||||||||| |||||||| |   | |
Db    3047 AGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGT 3106

Qy    3995 GGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCG 4054
           |||| ||||  |||| || |||||| || ||||| ||||| ||||||||||| ||||  |
Db    3107 GGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAG 3166

Qy    4055 TGAGGTGCTCTGGACAGTCGCTGAA------------------ATCACTGAATGCCTCCT 4096
           |||  ||| | |     || || |                   ||      |  ||| |
Db    3167 TGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCAT 3226

Qy    4097 CAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTA 4156
            | |||| | |  |||||  |  | | ||| ||||| |  ||||  |||    | ||
Db    3227 CCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCG 3286

Qy    4157 TTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAA 4216
            |   |||  ||   | |      |  |  | ||| | |||    ||  ||    ||||||
Db    3287 TCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACA---GCGGCTTGCAGTTTCCT 3343

Qy    4217 CCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGA 4276
            | |||  |    |   || |  |  |  |  | || ||  ||||||  |  ||||| ||
Db    3344 CAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATG 3403

Qy    4277 GAGAGGACCCACATGGGACAAGAACCTCAGATGAC 4311
            ||| || |     |   |   | |    |||||| | |
Db    3404 CAGATGATCTGGACCTAATGAATTCCTCAGGAGGC 3438
```

RESULT 10
AAI58650
ID   AAI58650 standard; cDNA; 3811 BP.
XX
AC   AAI58650;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 853.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39494.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 853; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3811 BP; 1018 A; 751 C; 1081 G; 961 T; 0 other;

Query Match 28.9%; Score 1339.2; DB 22; Length 3811;
Best Local Similarity 63.5%; Pred. No. 0;

```
Matches 2179;  Conservative    0;  Mismatches 1128;  Indels  123;  Gaps    4;

Qy   1002 GCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGG 1061
Db    107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

Qy   1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db    167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGTTCTCTTGGAGGAAC 226

Qy   1122 AGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAG 1181
Db    227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAAGTGAA 286

Qy   1182 AATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT 1241
Db    287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 346

Qy   1242 GGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAG 1301
Db    347 GATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAATTCCAG 406

Qy   1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCT 1361
Db    407 TGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCAGCTCT 466

Qy   1362 CTGGGACTGCACATATGATGGAAAAGCAAAGC---GAACATGCTTCCGAAGATCAGATGC 1418
Db    467 TTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAAGATGC 526

Qy   1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGGGCTCATAGCCC 1478
Db    527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586

Qy   1479 CTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG 1538
Db    587 GTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGATGATAA 646

Qy   1539 ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTATGCA 1598
Db    647 CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCTGTCAG 706

Qy   1599 TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC 1658
Db    707 TTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT 766

Qy   1659 TTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCATAA 1718
Db    767 ATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAGCATAA 826

Qy   1719 TTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCTGAG 1778
Db    827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGCCTGAG 886

Qy   1779 GCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db    887 ACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946

Qy   1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
Db    947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT 1006

Qy   1899 GGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAAA 1958
Db   1007 GGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066

Qy   1959 AATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAGGAA 2018
Db   1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126

Qy   2019 CAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTGA 2078
Db   1127 CCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy   2079 TGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGTTGA 2138
Db   1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTTGA 1246

Qy   2139 GGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACATTGC 2198
Db   1247 GGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAAGC 1306

Qy   2199 TGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA 2258
Db   1307 TGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGTA 1366

Qy   2259 TTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGAAGC 2318
Db   1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAAAC 1426

Qy   2319 CTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATATGGA 2378
Db   1427 TTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 1486

Qy   2379 AGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db   1487 AGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCCTG 1546

Qy   2439 CTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGATTT 2498
Db   1547 TTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606

Qy   2499 CTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATCTCT 2558
Db   1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCTAT 1666

Qy   2559 TTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCAGTG 2618
Db   1667 CCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTCCAGTG 1726

Qy   2619 TGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACACTTG 2678
Db   1727 TGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGAACTTG 1786

Qy   2679 TATCCACAGCAGAGAAGTTGGAGTTGTCTG------------------------------ 2708
Db   1787 TAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGTAAGACCCAGAAAACATCTTTAAT 1846

Qy   2709 ------------------------------------------------------------ 2708
Db   1847 TGGTTCTTATACTGTGAAAGGGACAGGGTTAGGGAGTCATAGCTGTCTTTTTCTAAAGCC 1906

Qy   2709 ---------TTCCCGATATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGG 2759
Db   1907 CTGTCTCCTTCCAGGATACACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGTGAGGG 1966

Qy   2760 GCAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCC 2819
Db   1967 CAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCTCTGTAACTCTCACTGGGACAT 2026

Qy   2820 AGAAGATGCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGG 2879
Db   2027 AGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACCCCAGG 2086

Qy   2880 AGGAAAATATATTGGAGAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGG 2939
Db   2087 AGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAGGCATATGTTTCACTGCACTGG 2146

Qy   2940 GAATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCA 2999
Db   2147 GACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCTAGGTGCTTCATTATGTCCTTC 2206

Qy   3000 TGGAAATACTGTCTCTGTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCT 3059
Db   2207 AGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTCCCAAACACTGTCCTCGTGCAA 2266
```

```
Qy   3060 CGCAAATGTATCTGACCCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTT 3119
Db   2267 TTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCCTGCAT 2326

Qy   3120 AGAGGACAAACGGCTCCGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGAT 3179
Db   2327 AGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTAGAGAT 2386

Qy   3180 CTATCACGACGGCTTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCCATGCCCA 3239
Db   2387 CTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAGTGATGCCCA 2446

Qy   3240 CGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTT 3299
Db   2447 CGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGGTTCTGCTCATTT 2506

Qy   3300 TGGGGAGGGGTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGGAACGGAGTCCCA 3359
Db   2507 TGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAAATGCAATGGAAAAGAATCCCG 2566

Qy   3360 CTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGC 3419
Db   2567 CATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAATTGCAGGCACAAGGAGGATGC 2626

Qy   3420 AGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAG 3479
Db   2627 GGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGACCAGTGAAGCCAGCAGAGAGGC 2686

Qy   3480 CTGTGCTGGGAGATTGGAAGTCTTCTATAACGGACCTGGGGCAGCGTCGGCAGGAGGAA 3539
Db   2687 CTGTGCAGGGCGTCTGGAAGTTTTTACAATGGAGCTTGGGGCACTGTTGGCAAGAGTAG 2746

Qy   3540 CATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGT 3599
Db   2747 CATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCAGACAAAGGGAA 2806

Qy   3600 TGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCA 3659
Db   2807 AATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTGGACAATGTTCA 2866

Qy   3660 GTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAAT 3719
Db   2867 GTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGGGAGAAGAGACT 2926

Qy   3720 CTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGG 3779
Db   2927 GGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAACAAGATAAGACTTCAGGAAGG 2986

Qy   3780 AGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTG 3839
Db   2987 ACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGAGGTTCCTGGGGGACAGTGTG 3046

Qy   3840 TGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTC 3899
Db   3047 TGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTTGGCTGTGGTCC 3106

Qy   3900 TGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGA 3959
Db   3107 AGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCGATATGGCTCAA 3166

Qy   3960 TGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGG 4019
Db   3167 TGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGATTGTCCTGCCAGACGCTGGGG 3226

Qy   4020 ACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAA 4079
Db   3227 CCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAATTGCACAGATATTTCAGTGCA 3286

Qy   4080 ------------------ATCACTGAATGCCTCCTCAGGTCATTTAGCACTTATTTTATC 4121
Db   3287 GAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCGTCAGTCATCCTTTATTGCAGT 3346

Qy   4122 CAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACGTGGTGCCGAGT 4181
Db   3347 CGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTCGCATTATTCTTCTTGACTAAAAA 3406

Qy   4182 TCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGTTCTCTCGAGGA 4241
Db   3407 GCGAAGACAGAGACA---GCGGCTTGCAGTTTCCTCAAGAGGAGAGAACTTAGTCCACCA 3463

Qy   4242 GAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACATGGGACAAGAAC 4301
Db   3464 AATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGATGATCTGGACCTAATGAATTC 3523

Qy   4302 CTCAGATGAC 4311
Db   3524 CTCAGGAGGC 3533
```

RESULT 11
AAF45141

```
ID   AAF45141 standard; cDNA; 4308 BP.
XX
AC   AAF45141;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Bovine WC1 ORF.
XX
KW   Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS   Bos sp.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
DR   P-PSDB; AAB66088.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Disclosure; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder. The present sequence is a
CC   sequence used in a sequence homology alignment with the TANGO/INTERCEPT
CC   sequences of the present invention.
XX
SQ   Sequence 4308 BP; 866 A; 1093 C; 1410 G; 939 T; 0 other;
```

```
Query Match             19.2%;  Score 886.8;  DB 22;  Length 4308;
Best Local Similarity   54.9%;  Pred. No. 1.8e-214;
Matches 2061;  Conservative    0;  Mismatches 1527;  Indels  168;  Gaps    9;

Qy    483 TCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATT 542
Db     75 TCTGGAGCTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCA 134

Qy    543 CCAAGAAAGGTGGGGACTATATGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGT 602
Db    135 CCAAGGAGAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGT 194

Qy    603 GTGCAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGC 662
Db    195 GTGCAGACAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTTGGGCC 254

Qy    663 TGTATTGCGCCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTG 722
Db    255 AGGACTTGGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTCAG 314

Qy    723 GAATTGCAGACATCGTGGATGGGGAAATCATGACTGCAGTCACAATGAGGATGTCACATT 782
Db    315 TGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGGGA 374

Qy    783 AACTTGTTATGATAGTAGTGATCTTGAACTAAGGCTTGTAGGTGGAACTAACCGCTGTAT 842
Db    375 TGCTGGAGTAGTCTGCTCAGGATTTG---TGCGTCTGGCTGGAGGGGATGGACCCTGCTC 431

Qy    843 GGGGAGAGTAGAGCTGAAAATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGGAA 902
Db    432 AGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGATGGGAACTTCAC 491

Qy    903 CAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACTTCGC 962
Db    492 ACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCTGTGTCTGTCCT 551

Qy    963 TGGCTTGCCTCATTTGCAGTCAGGGTCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTC 1022
Db    552 GGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAGTTCAGGTGTGA 611

Qy   1023 CGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCCGGAACCGTCAATTTTGACTGTCT 1082
Db    612 GGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGGGGCACGTGTCA 671

Qy   1083 TCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACTGCGACTAGC 1142
Db    672 CCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGA---AGTCCGGCTCATGAC 728

Qy   1143 AGATGGAAGTAACAATTGTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGAC 1202
Db    729 AAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGACAATGGAGAGC 788

Qy   1203 AATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATG 1262
Db    789 GCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGTCAGCTCGGCTG 848

Qy   1263 TCCGTTCAGCGTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTG 1322
Db    849 TGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGTGATCAGATCCT 908

Qy   1323 GATAAACAGCATATCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGG 1382
Db    909 AACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGACTGC 968

Qy   1383 AAAAGCAAAGCGAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGC 1442
Db    969 CCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAACCA 1028

Qy   1443 AGATCTGG---------------------------------------------------- 1450
Db   1029 GATCCAGGTGCTTCCCCAGTGCAACGACTCCGTGTCTCAACCTACAGGCTCTGCGGCCTC 1088

Qy   1451 ----------------------------------ACCTAAGGCTTGTCGGGGCTCATAG 1475
Db   1089 AGAGGACAGCGCCCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGGCGG 1148

Qy   1476 CCCCTGTTATGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTGTGTCATGA 1535
Db   1149 TCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGA 1208

Qy   1536 CAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGAAAGCCTAT 1595
Db   1209 CGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGCCCT 1268

Qy   1596 GCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGT 1655
Db   1269 CAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGACAACTT 1328

Qy   1656 TTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGGGGAAAGCA 1715
Db   1329 GAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCAGCA 1388

Qy   1716 TAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCCT 1775
Db   1389 CAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCAGAGT------TCCTGGCCCT 1442

Qy   1776 GAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACG 1835
Db   1443 CAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAATGGGAC 1502

Qy   1836 GTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCA 1895
Db   1503 CTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGCAGACA 1562

Qy   1896 GCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGG 1955
Db   1563 GCTTGGCTGTGGGGACAGTGGAACCCTCAACTCTTCTGTTGCTCTTAGAGAAGGTTTTAG 1622

Qy   1956 AAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAG 2015
Db   1623 GCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACCTCTCTCTGGCAGTGTCC 1682

Qy   2016 GAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTC 2075
Db   1683 TTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAAGCCTATATCTGGTG--- 1739

Qy   2076 TGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAAAGT 2135
Db   1740 ---TGCAGACAGCAGACAGATCCGCCTGGTGGATGGAGGTGGTCGCTGCTCTGGGAGAGT 1796

Qy   2136 TGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATGAACAT 2195
Db   1797 GGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGACCGCTGGGACCTGGACGA 1856

Qy   2196 TGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCC 2255
Db   1857 TGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTGGACGCCACTGTCTCTTC 1916

Qy   2256 TCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGAGGGGA 2315
Db   1917 CTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTGAACTGCAGAGGAGAGGA 1976

Qy   2316 AGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTAAATAT 2375
Db   1977 GTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACACAACTGCAATCATCAAGA 2036

Qy   2376 GGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGATATGCC 2435
Db   2037 AGATGCAGGAGTCATCTGCTCAGGATTTGTGC------GTCTGGCTGGAGGAGATGGACC 2090

Qy   2436 CTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGATTCTGA 2495
Db   2091 CTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACCCCAGTGTCTGATGGAAA 2150

Qy   2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGATGCCATATC 2555
```

```
Db   2151 CTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGATGTGGCAAGGCTGTGTC 2210

Qy   2556 TCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAGTTCCA 2615
Db   2211 TGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTCTGGGCTGAAGAGTTCAG 2270

Qy   2616 GTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAAGACAC 2675
Db   2271 GTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTGCCCTGTCCAGGAGGCAC 2330

Qy   2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735
Db   2331 ATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTACACAGAAGTCCAGCTTAT 2390

Qy   2736 G---AATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGGGG 2792
Db   2391 GAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAGATCTCTGGACGATGGAG 2450

Qy   2793 CTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTCAG 2852
Db   2451 AGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTGTCTGTCGTCAGCTCGG 2510

Qy   2853 CTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGTGT 2912
Db   2511 CTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTGGAAGGAGGTGATCAGAT 2570

Qy   2913 GTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATGAC 2972
Db   2571 CTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTGTGGAGTTGTCCTGTGAC 2630

Qy   2973 AGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGAAG 3032
Db   2631 TGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGCTCAGGAAA 2690

Qy   3033 CCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCAGT 3092
Db   2691 CCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAACCTGCAGGCTCTGCGGC 2750

Qy   3093 TCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGGGA 3152
Db   2751 CTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTCCGCCTGGTGGACGGGGG 2810

Qy   3153 CAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGA 3212
Db   2811 CGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGA 2870

Qy   3213 TGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGC 3272
Db   2871 TGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAGCTGGGCTGTGGAGAAGC 2930

Qy   3273 CTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGATGA 3332
Db   2931 CCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGCTGGACGA 2990

Qy   3333 CCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCA 3392
Db   2991 CCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGGCG 3050

Qy   3393 GCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAG 3452
Db   3051 GCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCAGAGTTCCTGGCCCTCAG 3110

Qy   3453 GCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGG 3512
Db   3111 G------ATGGTGAGCGAGGACCAGCAGTGTGCTGGGTGGCTGGAGGTTTTCTACAACGG 3164

Qy   3513 GACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAG 3572
Db   3165 GACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACTGTGTCCGTGATCTGCAG 3224

Qy   3573 GCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTC 3632
Db   3225 ACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTTGGTCTCAGGGAAGGTTC 3284

Qy   3633 TGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTG 3692
Db   3285 TAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGATACCTCTCTCTGGCAGTG 3344

Qy   3693 CCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATG 3752
Db   3345 TCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAGGAAGCCTACATCTCATG 3404

Qy   3753 TGAAGATAGAATAA---------------------------------------------- 3766
Db   3405 TGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGCACAGACAGAGAGAAGCT 3464

Qy   3767 --GAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTC 3824
Db   3465 CCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAGGTGTGGCACAACGGCTC 3524

Qy   3825 CTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCA 3884
Db   3525 CTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCTGAGGTGGTGTGTCAGCA 3584

Qy   3885 GCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGG 3944
Db   3585 GCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCATTTGGCCCTGGAAATGG 3644

Qy   3945 AACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCA 4004
Db   3645 GAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCCTCCCTGTGGGACTGTGT 3704

Qy   4005 CGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCTGGCGTGAGGTGCTC 4064
Db   3705 TGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGATGCTGGTGTGAGGTGCTC 3764

Qy   4065 TGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG 4100
Db   3765 TGGTGTAAGGACAACATTGCCCACGACCACAGCAGG 3800
```

RESULT 12
AAA43017
ID   AAA43017 standard; cDNA; 820 BP.
XX
AC   AAA43017;
XX
DT   21-AUG-2000 (first entry)
XX
DE   Human secreted expressed sequence tag SEQ ID NO:1757.
XX
KW   Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
KW   expressed sequence tag; EST; probe; chemotactic; proliferative;
KW   immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW   thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW   antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW   antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW   cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW   vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW   insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW   lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW   central nervous system disorder; Alzheimer's disease; stroke;
KW   Parkinson's disease; Huntington's disease; coagulation disorder;
KW   haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW   tumour; infection; depression; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200021990-A1.
XX
PD   20-APR-2000.
XX
PF   15-OCT-1999;   99WO-US24205.
XX
PR   15-OCT-1998;   98US-0104435.

```
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M;
XX
DR.  WPI; 2000-317937/27.
XX
PT   Isolated polynucleotides, and encoded proteins, comprising secreted
PT   expressed sequence tags (sESTs), useful for treating various disorders
PT   such as autoimmune, infectious, and central nervous system disorders -
XX
PS   Claim 1; Page 527; 618pp; English.
XX
CC   AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC   sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC   tissue sources. The sESTs can have a range of activities depending on
CC   the tissues they were isolated from. The activities include:
CC   chemotactic; proliferative; immunomodulatory; haematopoietic;
CC   chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC   cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC   antiasthmatic; vulnerary; antiulcer; osteopathic; neuroprotective;
CC   nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC   anticonvulsant; and antidepressant. The sESTs can be used for gene
CC   therapy and in vaccines. The sESTs are useful as probes for the
CC   identification and isolation of full-length cDNAs and genomic DNA
CC   molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC   are useful in assays for determining biological activity and raising
CC   antibodies. They may be useful for treatment of autoimmune disorders
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression and
CC   psoriasis. AAA43420 to AAA43425 represent linker variants which are given
CC   in the exemplification of the present invention.
XX
SQ   Sequence 820 BP; 195 A; 155 C; 233 G; 236 T; 1 other;

  Query Match             16.0%;  Score 740;  DB 21;  Length 820;
  Best Local Similarity   97.2%;  Pred. No. 1.5e-177;
  Matches  774;  Conservative    0;  Mismatches   20;  Indels    2;  Gaps    2;

Qy         26 TAATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGA 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 TAATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGA 83

Qy         86 ACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCA 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCA 143

Qy        146 GTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGA 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGA 203

Qy        206 CAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGT-GGAACACT 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        204 CAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGGAACACT 263

Qy        265 ACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 ACTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGT 323

Qy        325 TTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGA 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 TTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGA 383

Qy        385 AATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCAT 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 AATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCAT 443

Qy        445 GGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGAT 504
              |||||||| |||||||||||||||||  |   |   | ||||||||||||||||||||||
Db        444 GGAGAAGAAGTTGGTGTGAACTGTTAACGGTGAAGCCATCTGGGTTTGAGGCTAGTGGAT 503

Qy        505 GGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGACTATA 564
              |||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        504 GGAAAC-ACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGAACTATA 562

Qy        565 TGTGATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCA 624
              ||||||||||||||||||| |||||  |||||||  ||||||||||||||||||||||||
Db        563 TGTGATGATGGGTGGAACTTAAATACCCCTGCCGTCCTGTGCAGGCAACTAGGATGTCCA 622

Qy        625 TCTTCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTG 684
              |||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||
Db        623 TCTTCTTTTATTTCTTCTGGAGTTGCTAACAGCCCTGCTGTATTGCGCCCCATTTGGCTG 682

Qy        685 GATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGG 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GATGACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGG 742

Qy        745 GGAAATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGAT 804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 GGAAATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGAT 802

Qy        805 CTTGAACTAAGGCTTG 820
              |||||||  ||||| |
Db        803 CTTGAACGTAGGCTCG 818


RESULT 13
AAH99258
ID   AAH99258 standard; cDNA; 690 BP.
XX
AC   AAH99258;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein encoding cDNA sequence SEQ ID NO:93.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US35017.
XX
PR   23-DEC-1999;   99US-0471275.
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   P-PSDB; AAM25317.
XX
```

```
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS   Claim 1; Page 341; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and
CC   cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production. The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders.
XX
SQ   Sequence 690 BP; 163 A; 134 C; 207 G; 186 T; 0 other;

  Query Match             13.9%;  Score 642;  DB 22;  Length 690;
  Best Local Similarity   99.2%;  Pred. No. 1.1e-152;
  Matches  645;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   1930 CTGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGA 1989
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 CGGGGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGA 99

Qy   1990 GATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCAC 2049
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 GATGAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCAC 159

Qy   2050 AGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGT 2109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 AGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGT 219

Qy   2110 GGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTG 2169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 GGAAGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTG 279

Qy   2170 TGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGG 2229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 TGTGCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGG 339

Qy   2230 TCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATG 2289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    340 TCTGCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATG 399

Qy   2290 TCGAATTCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGG 2349
          || |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    400 TCCAATTCTGGCTGCGCTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGG 459

Qy   2350 AAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAG 2409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460 AAACAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAG 519

Qy   2410 CCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGAC 2469
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    520 CCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCACAC 579

Qy   2470 ACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGA 2529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    580 ACATGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGA 639

Qy   2530 GAATTAAATTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGG 2579
          |||||||| |||||||||||||||||||||||||||||||||||||||||
Db    640 GAATTAAACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGG 689


RESULT 14
AAV89272
ID   AAV89272 standard; cDNA; 608 BP.
XX
AC   AAV89272;
XX
DT   15-FEB-1999  (first entry)
XX
DE   EST clone CF118.
XX
KW   Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
KW   tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
KW   receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9845436-A2.
XX
PD   15-OCT-1998.
XX
PF   10-APR-1998;   98WO-US06955.
XX
PR   10-APR-1997;   97US-0838821.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Agostino MJ,  Jacobs K,  Lavallie ER,  McCoy JM,  Merberg D;
PI   Racie LA,  Spaulding V,  Treacy M;
XX
DR   WPI; 1999-070077/06.
XX
PT   New polynucleotides encoding human secreted proteins - derived from
PT   e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT   ovary, pituitary, retina and colon cDNA libraries.
XX
PS   Claim 1; Page 165; 618pp; English.
XX
CC   The present sequence represents a human expressed sequence tag (EST).
CC   The polynucleotide, which is a secreted EST, and the encoded protein
CC   are predicted to have useful biological activities which would make
CC   them suitable for treating, preventing or ameliorating medical
CC   conditions in humans and animals, although no supporting data is
CC   given. Suggested activities include nutritional activity, immune
CC   stimulating or suppressing activity, haematopoiesis regulating
CC   activity, tissue growth activity, activin/inhibin activity,
CC   chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC   activity, receptor/ligand activity, anti-inflammatory activity,
CC   cadherin/tumour invasion suppressor activity, tumour inhibition
CC   activity. The polynucleotide may also be useful for gene therapy.
XX
SQ   Sequence 608 BP; 145 A; 109 C; 176 G; 178 T; 0 other;

  Query Match             12.2%;  Score 563;  DB 20;  Length 608;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-132;
  Matches  574;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy     26 TAATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGA 85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     24 TAATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGA 83

Qy     86 ACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCA 145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     84 ACCTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCA 143

Qy    146 GTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGA 205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 GTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGA 203
```

```
Qy     206 CAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTA 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     204 CAGTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTA 263

Qy     266 CTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCA-TTTTCTTTCGCCATGTTTCGT 324
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     264 CTGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTTCTTTCGCCATGTTTCGT 323

Qy     325 TTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGA 384
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     324 TTTGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGA 383

Qy     385 AATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCAT 444
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 AATGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCAT 443

Qy     445 GGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGAT 504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     444 GGAGAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGAT 503

Qy     505 GGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGACTATA 564
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     504 GGAAACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGACTATA 563

Qy     565 TGTGATGATGGGTGGAACTTGAATACTGCTGCCGT 599
           |||||||||||||||||||||||||||||||||||
Db     564 TGTGATGATGGGTGGAACTTGAATACTGCTGCCGT 598
```

```
RESULT 15
AAH99326
ID   AAH99326 standard; cDNA; 562 BP.
XX
AC   AAH99326;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein encoding cDNA sequence SEQ ID NO:161.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US35017.
XX
PR   23-DEC-1999;   99US-0471275.
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   P-PSDB; AAM25385.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS   Claim 1; Page 369; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and
CC   cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production, The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders.
XX
SQ   Sequence 562 BP; 125 A; 131 C; 189 G; 117 T; 0 other;

  Query Match             11.9%;  Score 551.4;  DB 22;  Length 562;
  Best Local Similarity   99.8%;  Pred. No. 1e-129;
  Matches  552;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    3485 CTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCA 3544
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 CCGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCA 69

Qy    3545 CCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCA 3604
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 CCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCA 129

Qy    3605 GCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTC 3664
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTC 189

Qy    3665 CTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCA 3724
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 CTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCA 249

Qy    3725 GCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACA 3784
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 GCCCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACA 309

Qy    3785 CCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATG 3844
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 CCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATG 369

Qy    3845 ACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTC 3904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 ACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTC 429

Qy    3905 TGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACA 3964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 TGGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACA 489

Qy    3965 TGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGA 4024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 TGCGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGA 549

Qy    4025 GTGACTGTGGACA 4037
           |||||||||||||
```

```
Db    550 GTGACTGTGGACA 562

Search completed: May 12, 2003, 03:08:03
Job time : 683 secs
```

```
OM protein - protein search, using sw model

Run on:          May 12, 2003, 03:08:10 ; Search time 50 Seconds
                                          (without alignments)
                                          3872.263 Million cell updates/sec

Title:           US-09-759-130B-381
Perfect score:   8138
Sequence:        1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_101002:*
                 1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                 2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                 3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                 4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                 5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                 6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                 7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                 8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                 9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                 10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                 11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                 12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                 13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                 14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                 15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                 16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                 17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                 18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                 19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                 20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                 21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                 22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                 23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8138	100.0	1453	22	AAB66037	Human TANGO 234.
2	7913	97.2	1413	22	AAB66039	Human TANGO 234 ma
3	7426	91.3	1319	22	AAB66040	Human TANGO 234 ex
4	6058	74.4	1120	22	AAU00396	Human secreted pro
5	4540	55.8	821	22	AAU27709	Human full-length
6	3441.5	42.3	1151	23	AAU97585	Human CD163 recept
7	3429	42.1	1121	22	AAM39493	Human polypeptide
8	3429	42.1	1124	22	AAM41279	Human polypeptide
9	3429	42.1	1124	22	AAM41280	Human polypeptide
10	3429	42.1	1156	23	AAU97584	Human CD163 recept
11	3424	42.1	1116	23	AAU97582	Human CD163 recept
12	3396.5	41.7	1149	23	AAU97583	Human CD163 recept
13	3396.5	41.7	1154	22	AAM39494	Human polypeptide
14	3082	37.9	1436	22	AAB66088	Bovine WC1 protein
15	2435	29.9	1785	19	AAW64591	Human SRCR protein
16	1763.5	21.7	898	23	AAU09878	Novel human secret
17	1183	14.5	225	22	AAM25317	Human protein sequ
18	1048.5	12.9	1290	18	AAW07609	Rat von Ebner's gl
19	1033	12.7	186	22	AAM25385	Human protein sequ
20	971.5	11.9	552	22	AAE09447	Human sbg14862SPER
21	921	11.3	422	22	AAE09446	Human sbg14862SPER
22	822.5	10.1	822	20	AAW99087	Human serine prote
23	822.5	10.1	875	20	AAW83361	Human neurotrypsin
24	753	9.3	143	22	AAU27881	Human contig polyp
25	725	8.9	141	22	ABB11125	Human M160 precurs
26	719.5	8.8	761	20	AAW99088	Mouse serine prote
27	719.5	8.8	761	20	AAW83362	Mouse neurotrypsin
28	712.5	8.8	666	19	AAW64590	Human SRCR protein
29	691.5	8.5	347	19	AAW68200	Human scavenger re
30	691.5	8.5	347	20	AAY13369	Amino acid sequenc
31	691.5	8.5	347	22	AAB80237	Human PRO229 prote
32	691.5	8.5	347	23	ABB95434	Human angiogenesis
33	691.5	8.5	347	23	ABB84828	Human PRO229 prote
34	691.5	8.5	347	23	AAU83648	Human PRO protein,
35	690.5	8.5	347	19	AAW64537	Human liver cell c
36	634.5	7.8	757	21	AAB19127	Polypeptide isolat
37	634	7.8	111	22	ABB12184	Human Cys-rich sca
38	633.5	7.8	757	23	ABB07650	Human lysyl-oxidas
39	632.5	7.8	732	22	AAE11927	Human CG153 (or C5
40	629	7.7	754	21	AAB00078	Murine lysyl oxida
41	627.5	7.7	753	22	AAE11935	Human CG153 (or C5
42	625.5	7.7	743	23	AAG66060	Human lysyl oxidas
43	624.5	7.7	641	21	AAB12307	Human secreted pro
44	622.5	7.6	752	23	ABB07653	Human lysyl-oxidas
45	622.5	7.6	753	21	AAB00073	Human lysyl oxidas

ALIGNMENTS

```
RESULT 1
AAB66037
ID   AAB66037 standard; Protein; 1453 AA.
XX
AC   AAB66037;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human TANGO 234.
XX
KW   TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety.
XX
OS   Homo sapiens.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
DR   N-PSDB; AAF45123, AAF45124.
```

```
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 8; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence   1453 AA;

  Query Match             100.0%;  Score 8138;  DB 22;  Length 1453;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1453;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60

Qy       61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120

Qy      121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180

Qy      181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240

Qy      241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300

Qy      301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360

Qy      361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420

Qy      421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480

Qy      481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540

Qy      541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600

Qy      601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660

Qy      661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720

Qy      721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780

Qy      781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840

Qy      841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900

Qy      901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960

Qy      961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020

Qy     1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080

Qy     1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140

Qy     1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200

Qy     1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260

Qy     1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320

Qy     1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380

Qy     1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440

Qy     1441 SLLGVLPASEATK 1453
            |||||||||||||
Db     1441 SLLGVLPASEATK 1453


RESULT 2
AAB66039
ID   AAB66039 standard; Protein; 1413 AA.
XX
AC   AAB66039;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human TANGO 234 mature protein.
XX
KW   TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 8; Pages 276-281; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence   1413 AA;

  Query Match             97.2%;  Score 7913;  DB 22;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      41 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 60

Qy     101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 120

Qy     161 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 180

Qy     221 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 240

Qy     281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 300

Qy     341 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 360

Qy     401 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 420

Qy     461 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 480

Qy     521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 540

Qy     581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 600

Qy     641 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 660

Qy     701 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 720

Qy     761 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 780

Qy     821 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 840

Qy     881 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 900

Qy     941 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960

Qy    1001 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1020

Qy    1061 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1080

Qy    1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1140

Qy    1181 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1200

Qy    1241 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1260

Qy    1301 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHL 1360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHL 1320

Qy    1361 ALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDP 1420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 ALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDP 1380

Qy    1421 HGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1453
           |||||||||||||||||||||||||||||||||
Db    1381 HGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1413


RESULT 3
AAB66040
ID   AAB66040 standard; Protein; 1319 AA.
XX
AC   AAB66040;
XX
DT   30-MAR-2001  (first entry)
XX
```

```
DE   Human TANGO 234 extracellular domain.
XX
KW   TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety.
XX
OS   Homo sapiens.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 8; Pages 281-287; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence   1319 AA;

  Query Match             91.3%;  Score 7426;  DB 22;  Length 1319;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1319;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      41 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 60

Qy     101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 120

Qy     161 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 180

Qy     221 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 240

Qy     281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 300

Qy     341 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 360

Qy     401 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 420

Qy     461 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 480

Qy     521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 540

Qy     581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 600

Qy     641 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 660

Qy     701 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 720

Qy     761 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 780

Qy     821 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 840

Qy     881 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 900

Qy     941 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960

Qy    1001 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1020

Qy    1061 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1080

Qy    1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1140

Qy    1181 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1200

Qy    1241 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1260

Qy    1301 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1319


RESULT 4
AAU00396
ID   AAU00396 standard; Protein; 1120 AA.
XX
AC   AAU00396;
XX
```

```
DT   04-JUL-2001  (first entry)
XX
DE   Human secreted protein, POLY8.
XX
KW   Human secreted protein; therapeutic; diagnostic; human; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200119856-A2.
XX
PD   22-MAR-2001.
XX
PF   13-SEP-2000; 2000WO-US25106.
XX
PR   13-SEP-1999;   99US-0153629.
PR   16-SEP-1999;   99US-0154520.
PR   20-SEP-1999;   99US-0154762.
PR   13-OCT-1999;   99US-0159231.
PR   12-SEP-2000; 2000US-0659634.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Fernandes E,  Herrmann JL,  Liu X,  Yang M,  Boldog FL;
XX
DR   WPI; 2001-244781/25.
DR   N-PSDB; AAS01217.
XX
PT   New POLYX polypeptide useful for treating or preventing a POLYX
PT   associated disorder, e.g. cancer -
XX
PS   Claim 9; Page 25-29; 152pp; English.
XX
CC   The sequence represents the amino acid sequence of human secreted
CC   protein, POLY8. POLYX nucleic acids, polypeptides and antibodies to
CC   POLYX can be used for treating or preventing a POLYX associated disorder
CC   in a subject, preferably a human. These can be used in the manufacture of
CC   a medicament for treating a syndrome associated with a human disease
CC   selected from a POLYX-associated disorder, where the therapeutic is a
CC   POLYX polypeptide, a POLYX nucleotide or a POLYX antibody. They may also
CC   be used to screen for a modulator of activity, or latency, or
CC   predisposition to a POLYX associated disorder, e.g. cancer.
XX
SQ   Sequence   1120 AA;

  Query Match            74.4%;  Score 6058;  DB 22;  Length 1120;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1088;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        363 DGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 DGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA 89

Qy        423 KPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAH 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 KPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAH 149

Qy        483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 209

Qy        543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 269

Qy        603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 RWGTVCDDGWNGKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 329

Qy        663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 389

Qy        723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 449

Qy        783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 509

Qy        843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 569

Qy        903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 629

Qy        963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 689

Qy       1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 749

Qy       1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        750 FNATVSAHFGEGSGPIWLDDLNCTGMESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 809

Qy       1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 LYSETETGSCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 869

Qy       1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 929

Qy       1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 989

Qy       1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFLTW 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        990 WDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFLTW 1049

Qy       1383 CRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDTSL 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1050 CRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDTSL 1109

Qy       1443 LGVLPASEATK 1453
              |||||||||||
Db       1110 LGVLPASEATK 1120


RESULT 5
AAU27709
ID   AAU27709 standard; Protein; 821 AA.
XX
AC   AAU27709;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human full-length polypeptide sequence #34.
XX
KW   Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW   mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW   cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW   nervous system disorder; inflammatory disorder; cell differentiation;
KW   angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW   genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW   cytostatic; antirheumatic; antiarthritic; vulnerary; antiinflammatory;
KW   antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW   neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW   immunostimulant; analgesic; gene therapy.
XX
```

```
OS   Homo sapiens.
XX
PN   WO200164834-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04926.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
PR   17-JUN-2000; 2000US-0597707.
PR   14-JUL-2000; 2000US-0616807.
PR   19-SEP-2000; 2000US-0664641.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Asundi V,  Zhang J,  Zhao QA,  Ren F;
PI   Xue AJ,  Yang Y,  Wehrman T,  Wang J,  Ma Y,  Wang D,  Chen R,  Xu C;
PI   Drmanac R;
XX
DR   WPI; 2001-589862/66.
DR   N-PSDB; AAS44609.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis, treatment of
PT   cancer, neurological, inflammatory disorders and for use in arrays for
PT   detection  -
XX
PS   Claim 10; SEQ ID No 206; 153pp; English.
XX
CC   Sequences AAU27676-AAU28019 represent full-length polypeptides and
CC   contig polypeptides of the invention. The proteins and their associated
CC   DNA sequences are useful for the treatment, diagnosis and prevention of
CC   various types of disorder in a mammalian subject such as a human, dog,
CC   monkey, mouse, hamster or rat. The disorders include cancers such as
CC   leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
CC   multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC   diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC   disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC   chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC   Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC   disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC   bowel disease. The sequences exhibit activity relating to angiogenesis,
CC   cell proliferation, cell differentiation, stem cell growth factor,
CC   activin or inhibin. Therefore, they can be used to manipulate stem cells
CC   in culture to give rise to neuroepithelial cells that can be used to
CC   augment or replace cells damaged by illness, accidental damage or genetic
CC   disorders. The sequences may also be used for regeneration of bone,
CC   cartilage, tendons and ligaments and in tissue repair and burn healing.
CC   Note: Some sequences for this patent did not form part of the printed
CC   specification, but were obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   821 AA;

  Query Match             55.8%;  Score 4540;  DB 22;  Length 821;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  820;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        633 MGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRL 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRL 60

Qy        693 VGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHI 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHI 120

Qy        753 LMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKH 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKH 180

Qy        813 ADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLA 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLA 240

Qy        873 LCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPE 932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPE 300

Qy        933 DARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHG 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHG 360

Qy        993 NTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIY 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIY 420

Qy       1053 HDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHL 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        421 HDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGMESHL 480

Qy       1113 WQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNI 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 WQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNI 540

Qy       1173 TTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRIS 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRIS 600

Qy       1233 SPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSA 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 SPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSA 660

Qy       1293 LAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 720

Qy       1353 LNASSGHLALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEME 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LNASSGHLALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEME 780

Qy       1413 TCLKREDPHGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1453
              |||||||||||||||||||||||||||||||||||||||||
Db        781 TCLKREDPHGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 821


RESULT 6
AAU97585
ID   AAU97585 standard; Protein; 1151 AA.
XX
AC   AAU97585;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human CD163 receptor extracellular variant protein sequence.
XX
KW   Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
KW   CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
KW   human immunodeficiency virus; haemolysis; haematological condition;
KW   aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;
KW   sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;
KW   lymphoma; leucopenia; splenectomia; inflammation; infection; cancer;
KW   autoimmunity; immunodeficiency; CD163 receptor extracellular variant.
XX
OS   Homo sapiens.
XX
PN   WO200232941-A2.
XX
PD   25-APR-2002.
XX
PF   12-OCT-2001; 2001WO-DK00671.
XX
PR   16-OCT-2000; 2000DK-0001543.
PR   11-JAN-2001; 2001DK-0000039.
PR   22-FEB-2001; 2001US-270120P.
XX
```

```
PA   (PROT-) PROTEOPHARMA APS.
XX
PI   Moestrup S,  Moller HJ;
XX
DR   WPI; 2002-452380/48.
XX
PT   New haptoglobin-haemoglobin complex or its mimics or receptors, useful
PT   in gene therapy, particularly for treating an individual suffering from
PT   e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity
PT   -
XX
PS   Claim 1; Fig 5; 70pp; English.
XX
CC   The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)
CC   complex, or its part or mimic, being operably linked to a substance,
CC   capable of binding a CD163 receptor and/or a CD163 receptor variant.
CC   The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor
CC   are useful as a medicament or in the manufacture of a medicament for
CC   treating an individual. The medicament comprises at least one drug to be
CC   delivered to a cell expressing a CD163 receptor or a CD163 variant. The
CC   medicament may also comprise at least one gene to be delivered to a cell
CC   expressing a CD163 receptor or a CD163 variant. The medicament is an
CC   antibiotic or an anticancer drug, e.g. an anti-HIV (human
CC   immunodeficiency virus) drug. In particular, the Hp-Hb complex or its
CC   mimic, the CD163 variant or the CD163 receptor, or the medicament are
CC   useful for treating haemolysis and/or other haematological conditions,
CC   aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia,
CC   sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,
CC   lymphoma, leucopenia, splenectomia, inflammation, infection, cancer,
CC   autoimmunity or immunodeficiency. The method involves drug-delivery
CC   treatment or gene delivery treatment of an individual. The medicament is
CC   also useful for inhibiting (in vivo) uptake of haemoglobin in cells
CC   presenting a CD163 receptor. The CD163 variant is also useful in
CC   diagnosing, monitoring and/or controlling a condition in an individual.
CC   In particular, the CD163 variant is useful for identifying macrophages in
CC   a biological sample from an individual, for identifying at least one
CC   Hp-Hb complex in serum and/or plasma of an individual, or for removing at
CC   least one Hp-Hb complex in serum and/or plasma of an individual. The
CC   present amino acid sequence represents the human CD163 receptor
CC   extracellular variant of the invention.
XX
SQ   Sequence   1151 AA;

  Query Match             42.3%;  Score 3441.5;  DB 23;  Length 1151;
  Best Local Similarity   55.2%;  Pred. No. 2.7e-239;
  Matches  608;  Conservative  166;  Mismatches  311;  Indels   17;  Gaps    6;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db    581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db    641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db    701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db    761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db    821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGC 1434
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADDLDLMNSSEN--SHES 1115

Qy   1435 EDASDTSLLGV---LPASEATK 1453
Db   1116 ADFSAAELISVSKFLPISGMEK 1137


RESULT 7
AAM39493
ID   AAM39493 standard; Protein; 1121 AA.
XX
AC   AAM39493;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2638.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
```

```
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58649.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2638; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   1121 AA;

  Query Match            42.1%;  Score 3429;  DB 22;  Length 1121;
  Best Local Similarity  56.2%;  Pred. No. 2.1e-238;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels  12;  Gaps  4;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db     46 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 105

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db    106 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 165

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    166 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 225

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    226 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 285

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    286 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 345

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    346 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 405

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    406 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 465

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    466 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 525

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db    526 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 585

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db    586 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 645

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db    646 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 705

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db    706 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 765

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db    766 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 825

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db    826 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 885

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db    886 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 945

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db    946 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1005

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db   1006 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1065

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
Db   1066 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1107

RESULT 8
AAM41279
ID   AAM41279 standard; Protein; 1124 AA.
XX
AC   AAM41279;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 6210.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
```

```
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI60435.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 6210; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   1124 AA;

  Query Match             42.1%;  Score 3429;  DB 22;  Length 1124;
  Best Local Similarity   56.2%;  Pred. No. 2.1e-238;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels   12;  Gaps    4;

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
           | | |||| || | |||||||:: |:| |:|:  |  |    |:| |||| ::      |
Db      49 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 108

Qy     424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
            |  :  ||:: :|| ||||||||| :||  | : |  : |||| |||  ::|::||
Db     109 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 168

Qy     483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
           : | ||:|:|:|| ||||| | ::  :|:|:|:|:|| ||   :  |  :  | | ||||| ||
Db     169 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 228

Qy     543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
           : | |||| :|:|:|:| ||||||| | || | || |   |||| |   ||||||| |||
Db     229 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 288

Qy     603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
           |||:|||||:| ||| | || ||:::  :| ||| |:| |||| ||| | |  :| |
Db     289 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 348

Qy     663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
           ::  || : |:|:|| || ||| ||:|||| || ||||| ||| :|  :| :|  |||:
Db     349 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 408

Qy     723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
           |:|||||| ||||:: | : :   :   :   | | | | |||||  |:|    |
Db     409 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 468

Qy     783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
           || : |||||:||||| |:||||||||| ||| |:||||||| ||:||||| ||   :
Db     469 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 528

Qy     843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
           |:  | |||:|||  |||:||||| |:||:|||:    || || |||:||||||||::||
Db     529 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 588

Qy     903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
           ||||: |:|:||:  || ·||||::||| |||.|||:|| || |||| ||     |: : ::
Db     589 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 648

Qy     963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
           | | ||| | |   : :| :| ||| |        ||||:|: :| |  | ::  |   :
Db     649 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 708

Qy    1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
           || ||: |:|  :||||:| ||||||||||:| |||||| ||||||||||::|||| |
Db     709 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 768

Qy    1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
            ||| ||||||:||||||:: | | || :||| | |||| :|||||||||||||| :||
Db     769 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 828

Qy    1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
           | ||    |:|||||||||||| ||:||: :::     |:|||||| : | :: | | |  |
Db     829 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 888

Qy    1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
            ||||:||||    ::|||  |:|||:|::||:|||||||:::||:: | | ||||||||
Db     889 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 948

Qy    1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
           : |||||||||||| :|:||||||||| || | ::| |||||| |||:::::||||| |
Db     949 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1008

Qy    1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
           ||| |: || |:||||||| | |:  |::      :  ||   : |   | |:||| :|
Db    1009 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1068

Qy    1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
           :   |||  | |:|:  | ||:|  : :  :  || :|| :|
Db    1069 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1110


RESULT 9
AAM41280
ID   AAM41280 standard; Protein; 1124 AA.
XX
AC   AAM41280;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 6211.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
```

```
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI60436.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 6211; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   1124 AA;

  Query Match            42.1%;  Score 3429;  DB 22;  Length 1124;
  Best Local Similarity  56.2%;  Pred. No. 2.1e-238;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels   12;  Gaps    4;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db     49 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 108

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db    109 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 168

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    169 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 228

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    229 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 288

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    289 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 348

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    349 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 408

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    409 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 468

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    469 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 528

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db    529 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 588

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db    589 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 648

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db    649 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 708

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db    709 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 768

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db    769 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 828

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db    829 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 888

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db    889 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 948

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db    949 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1008

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db   1009 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1068

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
Db   1069 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1110


RESULT 10
AAU97584
ID   AAU97584 standard; Protein; 1156 AA.
XX
AC   AAU97584;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human CD163 receptor cytoplasmic variant 2 protein sequence.
XX
KW   Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
KW   CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
KW   human immunodeficiency virus; haemolysis; haematological condition;
KW   aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;
KW   sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;
KW   lymphoma; leucopenia; splenectomia; inflammation; infection; cancer;
KW   autoimmunity; immunodeficiency; CD163 receptor cytoplasmic variant 2.
XX
OS   Homo sapiens.
XX
PN   WO200232941-A2.
XX
PD   25-APR-2002.
XX
PF   12-OCT-2001; 2001WO-DK00671.
XX
PR   16-OCT-2000; 2000DK-0001543.
PR   11-JAN-2001; 2001DK-0000039.
```

```
PR   22-FEB-2001; 2001US-270120P.
XX
PA   (PROT-) PROTEOPHARMA APS.
XX
PI   Moestrup S,  Moller HJ;
XX
DR   WPI; 2002-452380/48.
XX
PT   New haptoglobin-haemoglobin complex or its mimics or receptors, useful
PT   in gene therapy, particularly for treating an individual suffering from
PT   e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity
PT   -
XX
PS   Claim 1; Fig 5; 70pp; English.
XX
CC   The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)
CC   complex, or its part or mimic, being operably linked to a substance,
CC   capable of binding a CD163 receptor and/or a CD163 receptor variant.
CC   The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor
CC   are useful as a medicament or in the manufacture of a medicament for
CC   treating an individual. The medicament comprises at least one drug to be
CC   delivered to a cell expressing a CD163 receptor or a CD163 variant. The
CC   medicament may also comprise at least one gene to be delivered to a cell
CC   expressing a CD163 receptor or a CD163 variant. The medicament is an
CC   antibiotic or an anticancer drug, e.g. an anti-HIV (human
CC   immunodeficiency virus) drug. In particular, the Hp-Hb complex or its
CC   mimic, the CD163 variant or the CD163 receptor, or the medicament are
CC   useful for treating haemolysis and/or other haematological conditions,
CC   aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia,
CC   sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,
CC   lymphoma, leucopenia, splenectomia, inflammation, infection, cancer,
CC   autoimmunity or immunodeficiency. The method involves drug-delivery
CC   treatment or gene delivery treatment of an individual. The medicament is
CC   also useful for inhibiting (in vivo) uptake of haemoglobin in cells
CC   presenting a CD163 receptor. The CD163 variant is also useful in
CC   diagnosing, monitoring and/or controlling a condition in an individual.
CC   In particular, the CD163 variant is useful for identifying macrophages in
CC   a biological sample from an individual, for identifying at least one
CC   Hp-Hb complex in serum and/or plasma of an individual, or for removing at
CC   least one Hp-Hb complex in serum and/or plasma of an individual. The
CC   present amino acid sequence represents the human CD163 receptor
CC   cytoplasmic variant 2 of the invention.
XX
SQ   Sequence   1156 AA;

  Query Match             42.1%;  Score 3429;  DB 23;  Length 1156;
  Best Local Similarity   56.2%;  Pred. No. 2.2e-238;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels   12;  Gaps    4;

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db      41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy     424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db     101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy     483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db     161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy     543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db     221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy     603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db     281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy     663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db     341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy     723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db     401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy     783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db     461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy     843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db     521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy     903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db     581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy     963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db     641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy    1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db     701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy    1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db     761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy    1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db     821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy    1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db     881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy    1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db     941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy    1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db    1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy    1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
Db    1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102


RESULT 11
AAU97582
ID   AAU97582 standard; Protein; 1116 AA.
XX
AC   AAU97582;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human CD163 receptor protein sequence.
XX
KW   Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
KW   CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
KW   human immunodeficiency virus; haemolysis; haematological condition;
KW   aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;
KW   sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;
KW   lymphoma; leucopenia; splenectomia; inflammation; infection; cancer;
KW   autoimmunity; immunodeficiency.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          4..146
FT                   /label= Domain_1
FT                   /note= "Specifically claimed in claim 6"
```

```
FT   Domain          154..253
FT                   /label= Domain_2
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          261..360
FT                   /label= Domain_3
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          368..467
FT                   /label= Domain_4
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          473..572
FT                   /label= Domain_5
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          578..677
FT                   /label= Domain_6
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          714..814
FT                   /label= Domain_7
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          819..920
FT                   /label= Domain_8
FT                   /note= "Specifically claimed in claim 6"
FT   Domain          924..1023
FT                   /label= Domain_9
FT                   /note= "Specifically claimed in claim 6"
XX
PN   WO200232941-A2.
XX
PD   25-APR-2002.
XX
PF   12-OCT-2001; 2001WO-DK00671.
XX
PR   16-OCT-2000; 2000DK-0001543.
PR   11-JAN-2001; 2001DK-0000039.
PR   22-FEB-2001; 2001US-270120P.
XX
PA   (PROT-) PROTEOPHARMA APS.
XX
PI   Moestrup S,  Moller HJ;
XX
DR   WPI; 2002-452380/48.
XX
PT   New haptoglobin-haemoglobin complex or its mimics or receptors, useful
PT   in gene therapy, particularly for treating an individual suffering from
PT   e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity
PT   -
XX
PS   Claim 1; Fig 5; 70pp; English.
XX
CC   The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)
CC   complex, or its part or mimic, being operably linked to a substance,
CC   capable of binding a CD163 receptor and/or a CD163 receptor variant.
CC   The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor
CC   are useful as a medicament or in the manufacture of a medicament for
CC   treating an individual. The medicament comprises at least one drug to be
CC   delivered to a cell expressing a CD163 receptor or a CD163 variant. The
CC   medicament may also comprise at least one gene to be delivered to a cell
CC   expressing a CD163 receptor or a CD163 variant. The medicament is an
CC   antibiotic or an anticancer drug, e.g. an anti-HIV (human
CC   immunodeficiency virus) drug. In particular, the Hp-Hb complex or its
CC   mimic, the CD163 variant or the CD163 receptor, or the medicament are
CC   useful for treating haemolysis and/or other haematological conditions,
CC   aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia,
CC   sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,
CC   lymphoma, leucopenia, splenectomia, inflammation, infection, cancer,
CC   autoimmunity or immunodeficiency. The method involves drug-delivery
CC   treatment or gene delivery treatment of an individual. The medicament is
CC   also useful for inhibiting (in vivo) uptake of haemoglobin in cells
CC   presenting a CD163 receptor. The CD163 variant is also useful in
CC   diagnosing, monitoring and/or controlling a condition in an individual.
CC   In particular, the CD163 variant is useful for identifying macrophages in
CC   a biological sample from an individual, for identifying at least one
CC   Hp-Hb complex in serum and/or plasma of an individual, or for removing at
CC   least one Hp-Hb complex in serum and/or plasma of an individual. The
CC   present amino acid sequence represents the human CD163 receptor of the
CC   invention.
XX
SQ   Sequence   1116 AA;

  Query Match             42.1%;  Score 3424;  DB 23;  Length 1116;
  Best Local Similarity   56.1%;  Pred. No. 4.8e-238;
  Matches  597;  Conservative  161;  Mismatches  295;  Indels   12;  Gaps    4;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
           | | |||| || |  |||||||:: |:| |:|:  |  |    |:| ||||| ::     |
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
            |  :  ||:: :||  ||||||||  :||    | : |   : |||| |||  ::|::||
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
           : | ||:|:|:|| ||||| |  ::   :|:|:|:|| ||    :     | : | | ||||| ||
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
           : | |||| :|:|:| ||||||| | ||    | ||   |     |||| |   ||||||| |||
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
           |||:||||||:|  |||  |  ||  ||:::   :|   |||  |:| ||||  |||  |  |    :|  |
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
           ::   ||  :  |:|:||  ||  |||  ||:||||  ||  |||||  |||  :|  :|  :|  |||:
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
           |:||||||  ||||::  |  :  :     :    :    :     |  |  |  |  |||||    |:|       |
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
            || :  |||||:|||||  |:||||||||||  |||  |:||||||| ||:||||||  ||   :
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
           |:    |  |||:|||    |||:|||||  |:||:|||:      ||  ||  |||:||||||||||::||
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
           ||||:  |:|:||:   ||   |||||:||:|||  |||  |||:||  ||  ||||  ||     |:  :  ::
Db    581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
            |  |  |||  |  |     :  :|  :|  :|  |||   |          ||||:|:  :|  |    |  ::     |        :
Db    641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
           || ||: |:|  :||||:|  ||||||||||:|  |||  |||  |||||||||||::||||  |
Db    701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTTCDDSWDLSDAHVVCRQLGCGEA 760

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
           |||  ||||||||:|||||::  |   |   ||  :|||   |  ||||  :|||||||||||||||  :||
Db    761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
           |  ||     |:||||||||||||  ||:||:  :::       ||:|||||||  :  |  ::  |  |  |   |
Db    821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
            ||||::||||     ::|||  |:|||:|::||:||||||||:::||::  |  |  ||||||||
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
           | |||||||||||||  :|:||||||||  ||  |  ::|  ||||||  |||::::||||||  |
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000
```

```
Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
          ||| |: || |:||||||| | |: |::       :   ||    : |   | |::|| :|
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
           :    |||  | |:|:   | ||:|    : :  : || :|| :|
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102


RESULT 12
AAU97583
ID   AAU97583 standard; Protein; 1149 AA.
XX
AC   AAU97583;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human CD163 receptor cytoplasmic variant 1 protein sequence.
XX
KW   Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
KW   CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
KW   human immunodeficiency virus; haemolysis; haematological condition;
KW   aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;
KW   sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;
KW   lymphoma; leucopenia; splenectomia; inflammation; infection; cancer;
KW   autoimmunity; immunodeficiency; CD163 receptor cytoplasmic variant 1.
XX
OS   Homo sapiens.
XX
PN   WO200232941-A2.
XX
PD   25-APR-2002.
XX
PF   12-OCT-2001; 2001WO-DK00671.
XX
PR   16-OCT-2000; 2000DK-0001543.
PR   11-JAN-2001; 2001DK-0000039.
PR   22-FEB-2001; 2001US-270120P.
XX
PA   (PROT-) PROTEOPHARMA APS.
XX
PI   Moestrup S,  Moller HJ;
XX
DR   WPI; 2002-452380/48.
XX
PT   New haptoglobin-haemoglobin complex or its mimics or receptors, useful
PT   in gene therapy, particularly for treating an individual suffering from
PT   e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity
PT    -
XX
PS   Claim 1; Fig 5; 70pp; English.
XX
CC   The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)
CC   complex, or its part or mimic, being operably linked to a substance,
CC   capable of binding a CD163 receptor and/or a CD163 receptor variant.
CC   The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor
CC   are useful as a medicament or in the manufacture of a medicament for
CC   treating an individual. The medicament comprises at least one drug to be
CC   delivered to a cell expressing a CD163 receptor or a CD163 variant. The
CC   medicament may also comprise at least one gene to be delivered to a cell
CC   expressing a CD163 receptor or a CD163 variant. The medicament is an
CC   antibiotic or an anticancer drug, e.g. an anti-HIV (human
CC   immunodeficiency virus) drug. In particular, the Hp-Hb complex or its
CC   mimic, the CD163 variant or the CD163 receptor, or the medicament are
CC   useful for treating haemolysis and/or other haematological conditions,
CC   aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia,
CC   sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,
CC   lymphoma, leucopenia, splenectomia, inflammation, infection, cancer,
CC   autoimmunity or immunodeficiency. The method involves drug-delivery
CC   treatment or gene delivery treatment of an individual. The medicament is
CC   also useful for inhibiting (in vivo) uptake of haemoglobin in cells
CC   presenting a CD163 receptor. The CD163 variant is also useful in
CC   diagnosing, monitoring and/or controlling a condition in an individual.
CC   In particular, the CD163 variant is useful for identifying macrophages in
CC   a biological sample from an individual, for identifying at least one
CC   Hp-Hb complex in serum and/or plasma of an individual, or for removing at
CC   least one Hp-Hb complex in serum and/or plasma of an individual. The
CC   present amino acid sequence represents the human CD163 receptor
CC   cytoplasmic variant 1 of the invention.
XX
SQ   Sequence   1149 AA;

  Query Match             41.7%;  Score 3396.5;  DB 23;  Length 1149;
  Best Local Similarity   54.4%;  Pred. No. 4.8e-236;
  Matches  597;  Conservative  161;  Mismatches  295;  Indels   45;  Gaps    5;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
          | | |||| || |  |||||||:: |:| |:|:  |  |     |:| ||||| ::       |
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
           |  :   ||:: :|| ||||||||| :||  | : |   : |||| ||| ::|::||
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
          : | ||:|:|:|| ||||| | ::  :|:|:|:|| ||   :     | : | | ||||| ||
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
          : | |||| :|:|:| |||||||  | ||   | ||   |   ||||  |    ||||||| |||
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
          |||:||||||:|  ||| |  || ||:::  :|   |||  |:| |||| |||  |  |   :| |
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
          ::  ||  :  |:|:|| || |||  ||:|||:|| || ||||| |||  |||  :|  :|  :| |||:
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
            |:||||||  ||||:: | :  :    :   :   :   |  |  |  |  |||||  |:|     |
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
            || : |||||:||||| |:|||||||||| ||| |:|||||||| ||:|||||| ||    :
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSR------ 896
          |:  | |||:|||  |||:||||| |:||:|||:    || || |||:||||||
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSSKTQKTS 580

Qy    897 --------------------------YTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHW 929
                                    ||::|||||: |:|:||:  || |||||::||
Db    581 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHW 640

Qy    930 DPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPC 989
           | ||| |||:|| || |||| ||       |: : ::| | ||| | |    : :| :| ||| |
Db    641 DIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLC 700

Qy    990 IHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRV 1049
                  ||||:|: :| |   | ::    |      :|| ||: |:|   :||||:|   ||||||
Db    701 PSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRV 760

Qy   1050 EIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTE 1109
          ||||:| ||||||| ||||||||||::|||| | ||| |||||||:||||||:: | | |
Db    761 EIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKE 820

Qy   1110 SHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGR 1169
           | :||| | |||| :||||||||||||||| :||| ||   |:||||||||||| ||:|||:
Db    821 SRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGK 880

Qy   1170 RNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWER 1229
           :::     |:||||||| :  | :: | | |   ||||::||||    ::||| |:|||:
Db    881 SSMSETTVGVVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEK 940
```

```
Qy   1230 RISSPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGC 1289
          |::||:|||||||:::||:: | | |||||||| |||||||||||| :|:|||||||
Db    941 RLASPSEETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGC 1000

Qy   1290 GSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQS 1349
           | || | ::| |||||| |||::::|||||| |||| |: || |:||||||| | |:  |
Db   1001 GPALKAFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDIS 1060

Qy   1350 LK------SLNASSGHLALILSSIFGLLLLVLFI--LFLTWCRVQKQKHLPLRVSTRRRG 1401
          ::    :   ||   : |   | |::|| :|:   ||| | |:|:  | ||:|
Db   1061 VQKTPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQR---LAVSSRGEN 1117

Qy   1402 SLEENLFHEMETCLKRED 1419
           : :  : || :|| :|
Db   1118 LVHQIQYREMNSCLNADD 1135
```

RESULT 13
AAM39494
ID   AAM39494 standard; Protein; 1154 AA.
XX
AC   AAM39494;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human polypeptide SEQ ID NO 2639.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58650.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2639; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   1154 AA;

```
  Query Match             41.7%;  Score 3396.5;  DB 22;  Length 1154;
  Best Local Similarity   54.4%;  Pred. No. 4.8e-236;
  Matches  597;  Conservative  161;  Mismatches  295;  Indels   45;  Gaps    5;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
          | | |||| || | ||||||||:: |:| |:|:  |  |    |:| |||| ::    |
Db     46 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 105

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
             |  :  ||:: :|| |||||||| :||   |  :  |   : |||| ||| ::|::||
Db    106 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 165

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
           : | ||:|:|:|:|| ||||| |  ::   :|:|:|:|:|| ||   :   | : | | ||||| ||
Db    166 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 225

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
          : | |||| :|:|:| ||||||| | ||  | ||  |    |||| |   ||||||| |||
Db    226 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 285

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
           |||:|||||:|  ||| | || ||::: :| ||| |:| |||| ||| | |  :| |
Db    286 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 345

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
          ::  || : |:|:|| || ||| ||:|||| || ||||| ||| :| :| :| |||:
Db    346 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 405

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
           |:||||||| ||||:: | : :    :  :  :  |  | | | |||| |:|    |
Db    406 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 465

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
           || : |||||:||||| |:|||||||||| ||| |:||||||| ||:|||||| ||  :
Db    466 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 525

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSR------ 896
          |:  | |||:|||  |||:|||||  |:||:|||:   || || |||:||||||
Db    526 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSSKTQKTS 585

Qy    897 --------------------------YTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHW 929
                                    ||::|||||||: |:|:||:  || ||||::||
Db    586 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHW 645

Qy    930 DPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPC 989
           | ||| |||:|| || |||| ||    |: : ::| | ||| | |    : :| :| ||| |
Db    646 DIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLC 705

Qy    990 IHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRV 1049
                  ||||:|: :| |  | ::  |    :|| ||: |:|  :||||:|  ||||||
Db    706 PSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRV 765

Qy   1050 EIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTE 1109
          ||||:| |||||||| ||||||||||::|||| | ||| ||||||:|||||||:: | | |
Db    766 EIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKE 825

Qy   1110 SHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGR 1169
          | :||| | |||| :|||||||||||||| :||| || |:||||||||||| ||:|:
Db    826 SRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGK 885

Qy   1170 RNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWER 1229
           :::    |:||||||| : | :: | | |  |  ||||::||||   ::||| |:|||:
```

```
Db     886 SSMSETTVGVVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEK 945

Qy    1230 RISSPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGC 1289
Db     946 RLASPSEETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGC 1005

Qy    1290 GSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQS 1349
Db    1006 GPALKAFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDIS 1065

Qy    1350 LK------SLNASSGHLALILSSIFGLLLLVLFI--LFLTWCRVQKQKHLPLRVSTRRRG 1401
Db    1066 VQKTPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQR---LAVSSRGEN 1122

Qy    1402 SLEENLFHEMETCLKRED 1419
Db    1123 LVHQIQYREMNSCLNADD 1140
```

RESULT 14
AAB66088
ID   AAB66088 standard; Protein; 1436 AA.
XX
AC   AAB66088;
XX
DT   30-MAR-2001 (first entry)
XX
DE   Bovine WC1 protein.
XX
KW   Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety.
XX
OS   Bos sp.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999; 99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX
DR   WPI; 2001-032313/04.
DR   N-PSDB; AAF45141.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Disclosure; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder. The present sequence is a
CC   sequence used in a sequence homology alignment with the TANGO/INTERCEPT
CC   sequences of the present invention.
XX
SQ   Sequence   1436 AA;

Query Match 37.9%; Score 3082; DB 22; Length 1436;
Best Local Similarity 41.7%; Pred. No. 3.1e-213;
Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;

```
Qy      28 CILLLNSCFLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCK 87
Db      13 CVLLLG-----TMVGGQALELRLKDGVHRCEGRVEVKHQGEWGTVDGYRWTLKDASVVCR 67

Qy      88 QLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQH---REWGSHNCYHGE 141
Db      68 QLGCGAAIGFPGGAYFGPGL---GPIWLLYTSCEGTESTVSDCEHSNIKDYRNDGYNHGR 124

Qy     142 DVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS 201
Db     125 DAGVVCSG----FVRLAGGDGPCSGRVEVHSGEAWIPVSDGNFTLATAQIICAELGCGKA 180

Qy     202 FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE 261
Db     181 VSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPCPGGTCHHSGSAQVVCSAYSEVR 240

Qy     262 LRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGS 321
Db     241 L-MTNGSSQCEGQVEMNISGQWRALCASHWSLANANVICRQLGCGVAISTPGGPHLVEEG 299

Qy     322 DVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGR 381
Db     300 DQILTARFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGN----------------- 342

Qy     382 VEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGN 441
Db     343 -QIQVLPQ---------------------CNDSV-----SQPTGSAAS----------- 363

Qy     442 ESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVC 501
Db     364 -------------------EDSAPYCSDSR--QLRLVDGGGPCAGRVEILDQGSWGTIC 401

Qy     502 HDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWG 561
Db     402 DDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKESHVWRCPSRGWG 461

Qy     562 KHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVC 621
Db     462 QHNCRHKQDAGVICS--EFLALRMVSEDQQCAGWLEVFYNGTWGSVCRNPMEDITVSTIC 519

Qy     622 SQLDCPSSIIGMGLGNAST----GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
Db     520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWNYNSCSPKEE 575

Qy     678 VGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAI 737
Db     576 AYIWCADS--RQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVCKQLGCGEAL 633

Qy     738 RVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRL 797
Db     634 DATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAGVICSGF--VRL 691

Qy     798 VGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLT 857
Db     692 AGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESDGQV 751

Qy     858 WAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLV-NGKSQCDGQVEIN 916
Db     752 WAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMKNGTSQCEGQVEMK 811

Qy     917 VLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLL 976
Db     812 ISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQISTAQFHCSGAESFL 871

Qy     977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036
```

```
Db     872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSAASEESSPYCSDSRQL 931

Qy    1037 RLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSG 1096
Db     932 RLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQLGCGEALNATGSAHFGAGSG 991

Qy    1097 PIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRL 1156
Db     992 PIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVICSEFLALRMVS--EDQQCAGWL 1049

Qy    1157 EVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTH 1216
Db    1050 EVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTSVGLREGSRPRWVDLIQCRKMD 1109

Qy    1217 ISIWQCLSAPWERRISSPAEETWITCE----------------DRIRVRGGDTECSGRVE 1260
Db    1110 TSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTAACTDREKLRLRGGDSECSGRVE 1169

Qy    1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
Db    1170 VWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEAVRSAAFGPGNGSIWLDEVQCGGRES 1229

Qy    1321 FLWDCHAKPWGQSDCGHKEDAGVRCSG----------------QSLKSLNASSGHLALIL 1364
Db    1230 SLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLPTTTAGTRTTSNSLPGIFSLPGVLCLIL 1289

Qy    1365 SSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTR 1424
Db    1290 GSLLFLVLVILVTQLLRW-RAER------RALSSYEDALAEAVYEELDYLLTQKEGLGSP 1342

Qy    1425 TS-DDTPNHGCEDASDTSLLGVLPASEATK 1453
Db    1343 DQMTDVPDENYDDAEEVPVPGTPSPSQGNE 1372
```

RESULT 15
AAW64591
ID   AAW64591 standard; Protein; 1785 AA.
XX
AC   AAW64591;
XX
DT   23-OCT-1998 (first entry)
XX
DE   Human SRCR protein.
XX
KW   Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
KW   nervous system; medullo-blastoma; glioma; breast; detection;
KW   autoantibody; ss.
XX
OS   Homo sapiens.
XX
PN   WO9830687-A2.
XX
PD   16-JUL-1998.
XX
PF   09-JAN-1998;   98WO-DE00096.
XX
PR   18-JUL-1997;   97DE-1030997.
PR   09-JAN-1997;   97DE-1000519.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI   Mollenhauer J,   Poustka A;
XX
DR   WPI; 1998-399136/34.
DR   N-PSDB; AAV49652.
XX
PT   Proteins containing scavenger receptor, cysteine rich domain -
PT   useful for diagnosis and treatment of tumours
XX
PS   Claim 2; Fig 2; 54pp; German.
XX
CC   This sequence represents a human protein which contains a SRCR (scavenger
CC   receptor, cysteine-rich) domain. The gene and encoded protein can be used
CC   to diagnose or treat tumours, particularly of the nervous system
CC   (medullo-blastoma or glioma) or breast. The DNA sequence and probes
CC   derived from it, are used to identify genes that express SRCR-domain
CC   containing proteins, to determine the form in which these proteins exist
CC   and to assess the significance of individual forms on cellular
CC   properties. The protein can be used to detect the presence of
CC   autoantibodies and antibodies which regulate its expression.
XX
SQ   Sequence 1785 AA;

```
  Query Match             29.9%;  Score 2435;  DB 19;  Length 1785;
  Best Local Similarity   36.1%;  Pred. No. 1.9e-166;
  Matches  533;  Conservative  155;  Mismatches  415;  Indels  374;  Gaps   28;

Qy      43 GTD--LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM--- 97
Db      95 GSDSGLALRLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGN 154

Qy      98 FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCY--------- 148
Db     155 AWFGQG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLR 211

Qy     149 ----------------GEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVV 192
Db     212 PESWPVRISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVV 271

Qy     193 CRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTL 252
Db     272 CRQLGCGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGV 331

Qy     253 TCY---------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHK 290
Db     332 ICSAPQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDS 391

Qy     291 WNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFD 350
Db     392 WDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHN 451

Qy     351 CLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQL 410
Db     452 CQHSEDAGVICSAA-----------------------HSWST------------------ 470

Qy     411 GCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470
Db     471 ------------PSPDTLPTITLPASTVGSESS--------------------------- 491

Qy     471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTY 530
Db     492 ---LALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVVCRQPGCGWAMSAPGNAR 548

Qy     531 FKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDA--------TW- 581
Db     549 FGQGSGPIVLDDVRCSGHESYPWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWP 608

Qy     582 ------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPS 628
Db     609 TSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGW 668

Qy     629 SIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICS----- 683
Db     669 AMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQ 728

Qy     684 -----------------DASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEV 726
Db     729 PTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANV 788

Qy     727 VCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEAS 786
Db     789 VCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAG 848

Qy     787 LICSAHR-QPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLS 845
```

```
Db    849 VICSASQSQP-------------TPSPDTWPT-------------------------- 867

Qy    846 VGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNG 905

Db    868 --------------SRASTAGSESTLAL--------------------------RLVNG 886

Qy    906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGH 965

Db    887 GDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQFGQGSGPIVLD 946

Qy    966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPYLSAVP 1023

Db    947 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGS 1006

Qy   1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAF 1083

Db   1007 ESSL-------ALRLVNGGDRCRGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAM 1059

Qy   1084 NATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRL 1143

Db   1060 SAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLTHNCGHHEDAGVICS---ATQI 1116

Qy   1144 YSET------------ETESCAGRLEVFY-NGTWGS---------------------- 1166

Db   1117 NSTTTDWWHPTTTTARPSSNCGGFL--FYASGTFSSPSYPAYYPNNAKCVWEIEVNSGY 1174

Qy   1167 ---VGRRNI--------------------TTAIAGIVC---RQLGCGENGVVSL---AP 1196

Db   1175 RINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICNDTRQIFTSSYNRMTIHFRSD 1234

Qy   1197 LSKTGSGFM-WVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTEC 1255

Db   1235 ISFQNTGFLAWYN----------------------SFPSDATL----RLVNLNSSYGLC 1267

Qy   1256 SGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRC 1315

Db   1268 AGRVEIYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVEC 1327

Qy   1316 KGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352

Db   1328 SGTESTLWQCRNRGWFSHNCNHREDAGVICSGNHLST 1364
```

Search completed: May 12, 2003, 06:37:17
Job time : 63 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

```
Run on:          May 12, 2003, 06:29:56 ; Search time 33 Seconds
                                          (without alignments)
                                          4232.828 Million cell updates/sec

Title:           US-09-759-130B-381
Perfect score:   8138
Sequence:        1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:      283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_73:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3441.5	42.3	1151	2	I38004	M130 antigen precu
2	3429	42.1	1156	2	I38005	M130 antigen precu
3	3396.5	41.7	1149	2	I38006	M130 antigen precu
4	3082	37.9	1436	2	A46496	antigen WC1.1 prec
5	2990	36.7	2403	2	A59386	sanko - human
6	2439.5	30.0	2153	2	T14893	scavenger receptor
7	2246.5	27.6	2043	2	T18524	scavenger receptor
8	2156.5	26.5	1594	2	T30549	hensin - rabbit
9	1962.5	24.1	2083	2	T42721	CRP-ductin-alpha p
10	1076.5	13.2	600	2	S56744	mucin (clone pGM7-
11	1048.5	12.9	1290	2	A57190	ebnerin precursor
12	866.5	10.6	1036	2	T17405	scavenger receptor
13	720.5	8.9	761	2	JC5759	brain-specific ser
14	665	8.2	918	2	JC4361	scavenger receptor
15	626.5	7.7	626	2	I49100	mscd6 precursor -
16	594.5	7.3	468	2	S26741	T-cell glycoprotei
17	520	6.4	532	2	A32751	speract receptor p
18	385	4.7	504	2	S56745	mucin (clone pGM31
19	314.5	3.9	577	2	A53202	cyclophilin C-asso
20	311.5	3.8	585	2	A47161	Mac-2-binding glyc
21	290.5	3.6	518	2	A55840	macrophage bacteri
22	290	3.6	458	2	B44407	macrophage scaveng
23	288	3.5	451	2	A38415	macrophage scaveng
24	279.5	3.4	453	2	S08276	macrophage scaveng
25	279.5	3.4	454	2	I46862	macrophage scaveng
26	273.5	3.4	454	2	A44407	macrophage scaveng
27	264	3.2	494	1	A29079	lymphocyte surface
28	258	3.2	5376	2	T42215	zonadhesin - mouse
29	254.5	3.1	495	1	A26396	T-cell surface gly
30	247.5	3.0	2471	2	A49128	cell-fate determin
31	246.5	3.0	2531	2	A46019	Notch-1 protein -
32	244.5	3.0	2871	2	A55624	fibrillin-1 precur
33	240	2.9	1574	2	T13954	MEGF6 protein - ra
34	239.5	2.9	4544	1	S02392	alpha-2-macroglobu
35	239	2.9	2531	2	S18188	notch protein homo
36	235	2.9	501	1	S43536	T-cell surface gly
37	234	2.9	2531	2	T31070	notch homolog - se
38	233	2.9	1700	2	S08167	Balbiani ring 3 pr
39	232	2.9	2352	2	T30201	Notch homolog prot
40	231.5	2.8	3002	2	A47221	fibrillin 1 precur
41	230	2.8	2703	1	A24420	notch protein - fr
42	230	2.8	2871	2	A55567	fibrillin I - bovi
43	224.5	2.8	4545	1	S25111	alpha-2-macroglobu
44	224	2.8	2907	2	A57278	fibrillin-2 precur
45	223.5	2.7	495	1	S11270	T-cell surface gly

ALIGNMENTS

```
RESULT 1
I38004
M130 antigen precursor, splice form 3 - human
N;Alternate names: M130 antigen cytosolic variant 1
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 17-Mar-2000
C;Accession: I38004; S36078
R;Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D
Eur. J. Immunol. 23, 2320-2325, 1993
A;Title: A new macrophage differentiation antigen which is a member of the scavenger
A;Reference number: I38003; MUID:93380506; PMID:8370408
A;Accession: I38004
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1151 <LAW>
A;Cross-references: EMBL:Z22969; NID:g312143; PIDN:CAA80542.1; PID:g312144
C;Comment: For other splice forms, see PIR:I34005 and PIR:I34006.
C;Genetics:
A;Gene: GDB:M130; GDB:MM130
A;Cross-references: GDB:119369; OMIM:133710; GDB:9958989
A;Map position: 10pter-10qter
A;Note: the locus has not been clearly established and there is some confusion in the
C;Superfamily: scavenger receptor cysteine-rich domain homology
C;Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protei
F;1-40/Domain: signal sequence #status predicted <SIG>
F;41-1151/Product: M130 antigen, splice form 3 #status predicted <MAT>
F;41-1045/Domain: extracellular #status predicted <EXT>
F;43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F;151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F;365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>
F;470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>
F;575-678/Domain: scavenger receptor cysteine-rich domain homology <SRC6>.
F;711-814/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F;816-920/Domain: scavenger receptor cysteine-rich domain homology <SRC8>
F;921-1024/Domain: scavenger receptor cysteine-rich domain homology <SRC9>
F;1046-1066/Domain: transmembrane #status predicted <TRN>
F;1067-1151/Domain: intracellular #status predicted <INT>
F;100,118,226,315,440,680,689,762,996/Binding site: carbohydrate (Asn) (covalent) #st

  Query Match             42.3%;  Score 3441.5;  DB 2;  Length 1151;
  Best Local Similarity   55.2%;  Pred. No. 2.1e-215;
  Matches  608;  Conservative  166;  Mismatches  311;  Indels   17;  Gaps    6;

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
           | | |||| || | |||||||:: |:| |:|:  |  |   |:| ||||| ::     |
Db      41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy     424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
            |  :  ||:: :|| ||||||||| :||  | : |    |||| ||| ::|::||
Db     101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160
```

```
Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db    581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db    641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db    701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db    761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db    821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGC 1434
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADDLDLMNSSEN--SHES 1115

Qy   1435 EDASDTSLLGV---LPASEATK 1453
Db   1116 ADFSAAELISVSKFLPISGMEK 1137
```

RESULT 2
I38005
M130 antigen precursor, splice form 4 - human
N;Alternate names: M130 antigen cytosolic variant 2

```
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 17-Mar-2000
C;Accession: I38005; S36079
R;Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D
Eur. J. Immunol. 23, 2320-2325, 1993
A;Title: A new macrophage differentiation antigen which is a member of the scavenger
A;Reference number: I38003; MUID:93380506; PMID:8370408
A;Accession: I38005
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1156 <LAW>
A;Cross-references: EMBL:Z22970; NID:g312145; PIDN:CAA80543.1; PID:g312146
C;Comment: For other splice forms, see PIR:I34004 and PIR:I34006.
C;Genetics:
A;Gene: GDB:M130; GDB:MM130
A;Cross-references: GDB:119369; OMIM:133710; GDB:9958989
A;Map position: 10pter-10qter
A;Note: the locus has not been clearly established and there is some confusion in the
C;Superfamily: scavenger receptor cysteine-rich domain homology
C;Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protei
F;1-40/Domain: signal sequence #status predicted <SIG>
F;41-1156/Product: M130 antigen, splice form 4 #status predicted <MAT>
F;41-1045/Domain: extracellular #status predicted <EXT>
F;43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F;151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F;365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>
F;470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>
F;575-678/Domain: scavenger receptor cysteine-rich domain homology <SRC6>
F;711-814/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F;816-920/Domain: scavenger receptor cysteine-rich domain homology <SRC8>
F;921-1024/Domain: scavenger receptor cysteine-rich domain homology <SRC9>
F;1046-1066/Domain: transmembrane #status predicted <TRN>
F;1067-1156/Domain: intracellular #status predicted <INT>
F;100,118,226,315,440,680,689,762,996/Binding site: carbohydrate (Asn) (covalent) #st

  Query Match             42.1%;  Score 3429;  DB 2;  Length 1156;
  Best Local Similarity   56.2%;  Pred. No. 1.3e-214;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels   12;  Gaps    4;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
```

```
             |:  | |||:|||  |||:|||||  |:||:|||:    || ||  |||:|||||||||::||
Db     521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy     903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
           ||||: |:|:||:  || |||||::||| ||| |||:|| || |||| ||     |: : ::
Db     581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy     963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
           | | ||| | |    : :| :| |||         ||||:|: :| :|  |   | :   :
Db     641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy    1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
           || ||: |:|  :||||:|  ||||||||||:| ||||||| |||||||||||::|||| |
Db     701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy    1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
            ||| ||||||:||||||:: | | || :||| | |||| :|||||||||||||| :||
Db     761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy    1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
           | ||     |:||||||||||| ||:||: :::       |:|||||| : | :: | | |  |
Db     821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy    1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
             ||||::||||    ::|||  |:|||:|::|:|||||||:||:|:  | | ||||||||
Db     881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy    1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
           | |||||||||||||| :|:||||||||| || |   ::| |||||| |||:::::|||||| |
Db     941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy    1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
           ||| |: || |:|||||||  | |:  |::       :   ||    :  |   | |::|| :|
Db    1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy    1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
           :    |||  | |:|:    | ||:|     : :   : || :||  :|
Db    1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102
```

RESULT 3
I38006
M130 antigen precursor, splice form 1 - human
N;Alternate names: M130 antigen, extracellular variant
N;Contains: M130 antigen, splice form 2
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-May-2000
C;Accession: I38006; I38003; S35768; S36077
R;Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D.Y.
Eur. J. Immunol. 23, 2320-2325, 1993
A;Title: A new macrophage differentiation antigen which is a member of the scavenger rec
A;Reference number: I38003; MUID:93380506; PMID:8370408
A;Accession: I38006
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1149 <LAW1>
A;Cross-references: EMBL:Z22971; NID:g312147; PIDN:CAA80544.1; PID:g312148
A;Accession: I38003
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-573,'R',608-1149 <LAW2>
A;Cross-references: EMBL:Z22968; NID:g312141; PIDN:CAA80541.1; PID:g312142
C;Comment: For other splice forms, see PIR:I34004 and PIR:I34005.
C;Genetics:
A;Gene: GDB:M130; GDB:MM130
A;Cross-references: GDB:119369; OMIM:133710; GDB:9958989
A;Map position: 10pter-10qter
A;Note: the locus has not been clearly established and there is some confusion in the as
C;Superfamily: scavenger receptor cysteine-rich domain homology
C;Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protein
F;1-40/Domain: signal sequence #status predicted <SIG>
F;41-1149/Product: M130 antigen, splice form 1 #status predicted <MAT1>
F;41-1078/Domain: extracellular #status predicted <EXT>
F;41-573,'R',608-1149/Product: M130 antigen, splice form 2 #status predicted <MAT2>
F;43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F;151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F;365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>
F;470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>
F;608-711/Domain: scavenger receptor cysteine-rich domain homology <SRC6>
F;744-847/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F;849-954/Domain: scavenger receptor cysteine-rich domain homology <SRC8>
F;955-1057/Domain: scavenger receptor cysteine-rich domain homology <SRC9>
F;1079-1099/Domain: transmembrane #status predicted <TRN>
F;1100-1149/Domain: intracellular #status predicted <INT>
F;100,118,226,315,440,713,722,795,1029/Binding site: carbohydrate (Asn) (covalent) #s

```
  Query Match             41.7%;  Score 3396.5;  DB 2;  Length 1149;
  Best Local Similarity   54.4%;  Pred. No. 1.7e-212;
  Matches  597;  Conservative  161;  Mismatches  295;  Indels   45;  Gaps    5;

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
           | | ||||  || |  | |||||||:: |:| |:|:   |   |     |:| |||||  ::       |
Db      41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy     424 FSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
             |  :   ||:: :|| |||||||| :||   | : |   : |||| ||| ::|::||
Db     101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy     483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
           : | ||:|:|:|| ||||| | ::  :|:|:|:||.||  :    | : | | ||||| ||
Db     161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy     543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
           : | |||| :|:|:| ||||||  | ||    | ||     ||||  |       |||||||  |||
Db     221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy     603 FWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
            |||:|||||:|  ||| | || ||:::  :|  |||  |:| |||| ||| |  | :| |
Db     281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy     663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
           ::  || : |:|:|| || ||| ||:|||| || ||||| ||| :|  :| :|  |||:
Db     341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy     723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
           |:|||||| ||||:: |  :   :   :   |  | |  | |  ||||| |:|     |
Db     401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy     783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
           || :  |||||:||||||| ||| |:|||||||| |||:||||||| ||:||||| ||: :
Db     461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy     843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSR------ 896
           |:  | |||:|||  |||:||||| |:||:|||:    || || |||:||||||
Db     521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSSKTQKTS 580

Qy     897 ---------------------------YTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHW 929
                                      ||::||||||: |:|:||:  || |||||::||
Db     581 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHW 640

Qy     930 DPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPC 989
           | ||| |||:|| |||| |||| ||      |: : ::| | ||| |  |    : :| :| |||  |
Db     641 DIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLC 700

Qy     990 IHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRV 1049
                ||||:|: :| |  |  :   |      :|| ||: |:|  :||||:| ||||||
Db     701 PSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRV 760

Qy    1050 EIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTE 1109
           ||||:| ||||||| |||||||||||::|||| | ||| ||||||:||||||:: | | |
Db     761 EIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKE 820

Qy    1110 SHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGR 1169
           | :||| | |||| :||||||||||||| :||  ||     |:||||||||||| ||:||:
Db     821 SRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGK 880
```

```
Qy    1170 RNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWER 1229
Db     881 SSMSETTVGVVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEK 940

Qy    1230 RISSPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGC 1289
Db     941 RLASPSEETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGC 1000

Qy    1290 GSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQS 1349
Db    1001 GPALKAFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDIS 1060

Qy    1350 LK------SLNASSGHLALILSSIFGLLLLVLFI--LFLTWCRVQKQKHLPLRVSTRRRG 1401
Db    1061 VQKTPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQR---LAVSSRGEN 1117

Qy    1402 SLEENLFHEMETCLKRED 1419
Db    1118 LVHQIQYREMNSCLNADD 1135
```

RESULT 4
A46496
antigen WC1.1 precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 18-Jun-1993 #sequence_revision 19-May-1994 #text_change 05-Nov-1999
C;Accession: A46496; I45834; S19913
R;Wijngaard, P.L.; Metzelaar, M.J.; MacHugh, N.D.; Morrison, W.I.; Clevers, H.C.
J. Immunol. 149, 3273-3277, 1992
A;Title: Molecular characterization of the WC1 antigen expressed specifically on bovine
A;Reference number: A46496; MUID:93056489; PMID:1431105
A;Accession: A46496
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1436 <WIJ>
A;Cross-references: EMBL:X63723; NID:g13; PIDN:CAA45255.1; PID:g14
A;Experimental source: CD4-CD8- gamma delta T lymphocytes
A;Note: sequence extracted from NCBI backbone (NCBIP:117475)
R;Wijngaard, P.L.; MacHugh, N.D.; Metzelaar, M.J.; Romberg, S.; Bensaid, A.; Pepin, L.;
J. Immunol. 152, 3476-3482, 1994
A;Title: Members of the novel WC1 gene family are differentially expressed on subsets of
A;Reference number: I45834; MUID:94194107; PMID:7511649
A;Accession: I45834
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1436 <WI2>
A;Cross-references: EMBL:X63723; NID:g13; PIDN:CAA45255.1; PID:g14
C;Genetics:
A;Gene: WC1.1
C;Superfamily: scavenger receptor cysteine-rich domain homology
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-1436/Product: WC1 antigen #status predicted <MAT>
F;25-131/Domain: scavenger receptor cysteine-rich domain homology <SR01>
F;132-234/Domain: scavenger receptor cysteine-rich domain homology <SR02>
F;236-340/Domain: scavenger receptor cysteine-rich domain homology <SR03>
F;373-476/Domain: scavenger receptor cysteine-rich domain homology <SR04>
F;478-581/Domain: scavenger receptor cysteine-rich domain homology <SR05>
F;583-686/Domain: scavenger receptor cysteine-rich domain homology <SR06>
F;687-789/Domain: scavenger receptor cysteine-rich domain homology <SR07>
F;791-895/Domain: scavenger receptor cysteine-rich domain homology <SR08>
F;928-1031/Domain: scavenger receptor cysteine-rich domain homology <SR09>
F;1033-1136/Domain: scavenger receptor cysteine-rich domain homology <SR10>
F;1152-1255/Domain: scavenger receptor cysteine-rich domain homology <SR11>

```
  Query Match             37.9%;  Score 3082;  DB 2;  Length 1436;
  Best Local Similarity   41.7%;  Pred. No. 5.6e-192;
  Matches  613;  Conservative  208;  Mismatches  495;  Indels  154;  Gaps   23;

Qy      28 CILLLNSCFLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCK 87
Db      13 CVLLLG-----TMVGGQALELRLKDGVHRCEGRVEVKHQGEWGTVDGYRWTLKDASVVCR 67

Qy      88 QLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQH---REWGSHNCYHGE 141
Db      68 QLGCGAAIGFPGGAYFGPGL---GPIWLLYTSCEGTESTVSDCEHSNIKDYRNDGYNHGR 124

Qy     142 DVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS 201
Db     125 DAGVVCSG----FVRLAGGDGPCSGRVEVHSGEAWIPVSDGNFTLATAQIICAELGCGKA 180

Qy     202 FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE 261
Db     181 VSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPCPGGTCHHSGSAQVVCSAYSEVR 240

Qy     262 LRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGS 321
Db     241 L-MTNGSSQCEGQVEMNISGQWRALCASHWSLANANVICRQLGCGVAISTPGGPHLVEEG 299

Qy     322 DVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGR 381
Db     300 DQILTARFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGN---------------- 342

Qy     382 VEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGN 441
Db     343 -QIQVLPQ---------------------CNDSV-----SQPTGSAAS---------- 363

Qy     442 ESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVC 501
Db     364 -------------------EDSAPYCSDSR--QLRLVDGGGPCAGRVEILDQGSWGTIC 401

Qy     502 HDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWG 561
Db     402 DDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKESHVWRCPSRGWG 461

Qy     562 KHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVC 621
Db     462 QHNCRHKQDAGVICS--EFLALRMVSEDQQCAGWLEVFYNGTWGSVCRNPMEDITVSTIC 519

Qy     622 SQLDCPSSIIGMGLGNAST----GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
Db     520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWNYNSCSPKEE 575

Qy     678 VGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAI 737
Db     576 AYIWCADS--RQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVCKQLGCGEAL 633

Qy     738 RVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRL 797
Db     634 DATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAGVICSGF--VRL 691

Qy     798 VGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLT 857
Db     692 AGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESDGQV 751

Qy     858 WAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLV-NGKSQCDGQVEIN 916
Db     752 WAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMKNGTSQCEGQVEMK 811

Qy     917 VLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLL 976
Db     812 ISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQISTAQFHCSGAESFL 871

Qy     977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036
Db     872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSAASEESSPYCSDSRQL 931

Qy    1037 RLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSG 1096
Db     932 RLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQLGCGEALNATGSAHFGAGSG 991

Qy    1097 PIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRL 1156
Db     992 PIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVICSEFLALRMVS--EDQQCAGWL 1049

Qy    1157 EVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTH 1216
```

```
Db    1050 EVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTSVGLREGSRPRWVDLIQCRKMD 1109

Qy    1217 ISIWQCLSAPWERRISSPAEETWITCE----------------DRIRVRGGDTECSGRVE 1260

Db    1110 TSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTAACTDREKLRLRGGDSECSGRVE 1169

Qy    1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320

Db    1170 VWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEAVRSAAFGPGNGSIWLDEVQCGGRES 1229

Qy    1321 FLWDCHAKPWGQSDCGHKEDAGVRCSG----------------QSLKSLNASSGHLALIL 1364

Db    1230 SLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLPTTTAGTRTTSNSLPGIFSLPGVLCLIL 1289

Qy    1365 SSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTR 1424

Db    1290 GSLLFLVLVILVTQLLRW-RAER------RALSSYEDALAEAVYEELDYLLTQKEGLGSP 1342

Qy    1425 TS-DDTPNHGCEDASDTSLLGVLPASEATK 1453

Db    1343 DQMTDVPDENYDDAEEVPVPGTPSPSQGNE 1372
```

RESULT 5
A59386
sanko - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 31-Dec-2001
C;Accession: A59386
R;sanko, S.
submitted to the Protein Sequence Database, March 2001
A;Reference number: A59386
A;Accession: A59386
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-2403 <SAN>

```
  Query Match             36.7%;  Score 2990;  DB 2;  Length 2403;
  Best Local Similarity   40.5%;  Pred. No. 9.2e-186;
  Matches  619;  Conservative  177;  Mismatches  504;  Indels  230;  Gaps   23;

Qy      44 TDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM---FRF 100

Db     230 SSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQF 289

Qy     101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG----------- 149

Db     290 GQG---SGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGVICSAPLSRPTPSPDT 346

Qy     150 -----------EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGC 198

Db     347 WPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDANVVCRQLGC 406

Qy     199 PSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDS- 257

Db     407 GWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSEDAGVICSDTL 466

Qy     258 -------------SDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLG 304

Db     467 PTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLG 526

Qy     305 CGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS-D 363

Db     527 CGWAMLAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHNGWLSHNCGHSEDAGVICSGP 586

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423

Db     587 ESSLALGLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWATSAPGNAR 646

Qy     424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICS-------------- 469

Db     647 FGQGSGPIVLDDVRCSGHESYLWSCPNNGWLSHNCGHHEDAGVICSAAQSRSTPRPDTLS 706

Qy     470 ----------DKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGC 519

Db     707 TITLPPSTVGSESSLTLRLVNGSDRCQGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGC 766

Qy     520 GKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCS--- 576

Db     767 GWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSVSQ 826

Qy     577 -----GDATW--------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAA 617

Db     827 SRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTSDA 886

Qy     618 AVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677

Db     887 NVVCRRLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHNCQHSED 946

Qy     678 VGVICSDA-----------------------SDMELRLVGGSSRCAGKVEVNVQGAVGI 713

Db     947 AGVICSAAHSWSTPSPDTLPTITLPASTVGSESSLALRLVNGGDRCQGRVEVLYQGSWGT 1006

Qy     714 LCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWE 773

Db    1007 VCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNG 1066

Qy     774 WKQTACHLNMEASLICSAHR-QP-----------------------RLVGADMPCSGRVE 809

Db    1067 WLSHNCGHSEDAGVICSASQSRPTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVE 1126

Qy     810 VKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSET 869

Db    1127 VLYRGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGHES 1186

Qy     870 HLALCPIVQHPEDTCIHSREVGVVCSRYTD-----------------------VRLVNG 905

Db    1187 YLWSCPHDGWLSHNCGHHEDAGVICSASQSQPTPSPDTWPTSHASTAGSESSLALRLVNG 1246

Qy     906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGH 965

Db    1247 GDRCQGRVEVLYRGPWGTVCDDYWDTNDANVVCRQLGCGWATSAPGNARFGQGSGPIVLD 1306

Qy     966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEG 1025

Db    1307 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPT-PSPDTWPTSHASTAGSE 1365

Qy    1026 SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNA 1085

Db    1366 SSL------ALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGWATSA 1419

Qy    1086 TVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC---------- 1135

Db    1420 PGSARFGQGSGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTPS 1479

Qy    1136 -------------SEFT-ALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVC 1181

Db    1480 PDTWPTSRASTAGSESTLALRLVN--GGDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVC 1537

Qy    1182 RQLGCGENGVVSLAPLSK---TGSGFMWVDDIQCPKTHISIWQCLSAPW----------- 1227

Db    1538 RQLGC---GWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDA 1594

Qy    1228 ------ERRISSPAEETWIT---------CEDRIRVR---GGDTECSGRVEIWHAGSWGT 1269

Db    1595 GVICSAAQSQSTPRPDTWLTTNLPALTVGSESSLALRLVNGGD-RCRGRVEVLYRGSWGT 1653

Qy    1270 VCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKP 1329

Db    1654 VCDDSWDTNDANVVCRQLGCGWAMSAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKG 1713

Qy    1330 WGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359

Db    1714 WLTHNCGHHEDAGVICSATQINSTTTDWWH 1743
```

```
RESULT 6
T14893
scavenger receptor cysteine-rich protein precursor - sea urchin (Strongylocentrotus purp
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: T14893
R;Pancer, Z.; Rast, J.P.; Davidson, E.H.
Immunogenetics 49, 773-786, 1999
A;Title: Origins of immunity: transcription factors and homologs of effector genes of th
A;Reference number: Z18253; MUID:99328904; PMID:10398804
A;Accession: T14893
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2153 <PAN>
A;Cross-references: EMBL:AF064259; NID:g4165052; PID:g4165053; PIDN:AAD08654.1
C;Genetics:
A;Gene: SRCR12

  Query Match             30.0%;  Score 2439.5;  DB 2;  Length 2153;
  Best Local Similarity   38.2%;  Pred. No. 4.3e-150;
  Matches  527;  Conservative  197;  Mismatches  560;  Indels   97;  Gaps   22;

Qy       32 LNSCFLISSFN-------------GTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWN 78
Db      113 LSSCAYVSSHNCAHSEDAGVVCAGSGDGNIRLVGGSHSSEGRVELYVNNQWGTVCDDLWD 172

Qy       79 TTASTVVCKQLGC-PFSFAMFRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNC 137
Db      173 LNDAQVACRQLGLGPAVSADAHFGEG---SGSILLDNVACIGTETSLLSCSHPGIESHNC 229

Qy      138 YHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLG 197
Db      230 GHSEDAGVICSSRFGSDIRVVGGSIPTEGRVEVFVNGAWGTVCDDLWDINDASVACRQLG 289

Qy      198 CPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDS 257
Db      290 FGRAISASGGASYGQGSGSIFLDNLACTGAESNLLSCPHNGVGSHNCGHGEDAGVLCSSS 349

Qy      258 SDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHL 317
Db      350 LSSDIRVVGGRSPTEGRVEVFVNRAWGTVCDDLWDVNDTNVACRQLGFGRAISAPGGASY 409

Qy      318 QSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGA-DLELRLADGSN 376
Db      410 GQGSGSILLDNLACTGAESSLLSCPHNGVGSHNCAHGEDAGVSCAPSSQESRVRLVGGLN 469

Qy      377 NCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKP--SNEARDIWIN 434
Db      470 NREGRVEIFLNNQWGTVCDDDWGTPDANVVCRQLGYPSG--GSARSSAYFGRGSVPILLD 527

Qy      435 SISCTGNESALWDCTYDGKAKRTCFRRSDAGVIC---SDKADLDLRLVGAHSPCYGRLEV 491
Db      528 NVGCSGNERSLELCSNNGIGVHNCGHQEDASVVCTGVTTMPGMNIRIVGGGSPSEGRVEV 587

Qy      492 KYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESN 551
Db      588 LVGHRWGTVCDDLWDINDANVVCRELGYSAATSATSSASFGQGSGDILLDDLRCSGTESS 647

Qy      552 IWDCEHSGWGKHNCVHREDVIVTCSGDATWG------LRLVGGSNRCSGRLEVYFQGRWG 605
Db      648 LLTCPHRGVGVHNCAHSEDAGVVCA-SITSGPVVGTYMRLVGGQNSRQGRLEISINNQWG 706

Qy      606 TVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNS 665
Db      707 TVCDDSWDINDATVVCRQLGFSSAVSAPTSAHFGQGSGTIWLDDVSCAGNENSLMDCGHR 766

Qy      666 GWGNNDCSHSEDVGVICSDASD--MELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNI 723
Db      767 GLGVHNCAHAEDAGVVCI-ASDGPLNIRLAGGRSGMEGRVEISLGGDWGTVCDDSWGIED 825

Qy      724 AEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNM 783
Db      826 AHVVCRQLGFGPALSAVTAASFGQGSGSILMDNVQCSGDEATIAECSHNGIGIHNCGHQE 885

Qy      784 EASLICS-AHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db      886 DAGVVCSRATGEVRLVGNSRQNEGRVEILISGRWGTVCDDLWDLRDADVICRQLGYGNAI 945

Qy      843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP---EDTCIHSREVGVVCSRYTD 899
Db      946 SAPHSSFFGPGRGDILLDDVSCTGSEDDILDC---SHPPIGTNNCGHSEDAGVVCD--LN 1000

Qy      900 VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERS 959
Db     1001 VRIVNGSRSNEGRVEILHDGSWATVCDDNWDIDDATVVCRQLGFPSANYAVAHAVFGQGS 1060

Qy      960 VRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYL 1019
Db     1061 GEIVLDDVECTGDEVSLIECQHAGLGTNNCGHSEDAGVICS-----------VNV----- 1104

Qy     1020 SAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGC 1079
Db     1105 -----------------RLADGNSPAEGRVEVYFDGQWGTVCDDNWGLQNGHVICRAVGF 1147

Qy     1080 GVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFT 1139
Db     1148 GKAIAVTDRARFGRGQGPIYLDNINCTGSETSLLNCSHSGLGTHDCDHSEDAGVICSQPD 1207

Qy     1140 ALRL---YSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAP 1196
Db     1208 TIRLAGGRSKYE-----GRVEILQNGAWGTVCDNQWDVHDAAVVCRELGYATTLEATSQA 1262

Qy     1197 LSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVR--GGDTE 1254
Db     1263 SFGQGTGAILLDDLRCSGREIRLVDCPHNGLGQHNCQHLEDAGVVCQDVIELRLVDGQTP 1322

Qy     1255 CSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMR 1314
Db     1323 NAGRVELRYQGQWGTVCDNGWDVRDAAVVCRELGFPRAIAANNQAFYGPGIGTVLLSDIS 1382

Qy     1315 CKGNESFLWDCHAKPWGQSDCGHKEDAGVRC------SGQSLKSLNASSGHLALILSSIF 1368
Db     1383 CGGQEATLFSCPNSGVGLNSCNHGEDAGVVCMPKVRLSG----SNNPNEGLVEVFMNGLW 1438

Qy     1369 G 1369
Db     1439 G 1439


RESULT 7
T18524
scavenger receptor cysteine-rich protein homolog srcrm2 - Geodia cydonium
C;Species: Geodia cydonium
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C;Accession: T18524
R;Blumbach, B.; Pancer, Z.; Diehl-Seifert, B.; Steffen, R.; Muenkner, J.; Mueller, I.
J. Cell Sci. 111, 2635-2644, 1998
A;Title: The putative sponge aggregation receptor: Isolation and characterisation of
A;Reference number: Z18947; MUID:98369060; PMID:9701562
A;Accession: T18524
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2043 <BLU>
A;Cross-references: EMBL:Y14953; NID:e1364818; PID:e1364819; PIDN:CAA75175.1
C;Genetics:
A;Gene: SRCRM2

  Query Match             27.6%;  Score 2246.5;  DB 2;  Length 2043;
  Best Local Similarity   32.4%;  Pred. No. 1.4e-137;
  Matches  514;  Conservative  196;  Mismatches  588;  Indels  287;  Gaps   31;

Qy        7 SWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLV-NGDGPCSGTVEVKF 65
Db      114 AWH-SVGSECSHQNNSGVVCT----------------DDSAVRLVEGGDMPNRGRVEVYH 156
```

```
Qy     66 QGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGNESALW 125
Db    157 ADEWGSICDDYFGFGEAVVICHELGFPSAEDSHGGGTFPQGDGLIWMDNLQCTGHESSIA 216

Qy    126 ECQHREWGSHNCYHGEDVGVNCYGEAN-----LGLRLVDGNNSCS-------GRVEVKFQ 173
Db    217 ECTFPGWGEHNCQHREDAGVTCTPATHDPFTYIPIRLNNGTTSSNATHGLQYGRVEVFIN 276

Qy    174 ERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLR---PIWLDDILCQGNELA 230
Db    277 NTWGTVCDDGWGIEDATIACRQLGFYGALLASSRSEYAQVAPSSVPIYWDEVLCHGDEVA 336

Qy    231 LWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE-----LRLVGGTNRCM--------GRVEL 277
Db    337 LSECHHRTTGVHDCSHYEDAGVFC-TTNDIQPAAPKVRLVRNNNGSVDEDGHVFEGRVEV 395

Qy    278 KIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESF 337
Db    396 NFYGVWGTICDDYWSLADGDVICKMLGFTYADEVGYSGAFGIGTGPIWMDDVQCDGTETR 455

Qy    338 LWDCRHSGTVNFDCLHQNDVSVICS---DGADLELRLADGSNNCSGRVEVRIHEQWWTIC 394
Db    456 IDHCDFPGWGVHNCYHFEDASVVCSYRQPGSD--VRLVGGSSSMEGRVEVLHNGVWGTVC 513

Qy    395 DQNWKNEQALVVCKQLGCPFSVFGSRRAKP------SNEARDIWINSISCTGNESALWDC 448
Db    514 DDSWGAEDAQVICHQLG----FYGTATAVRGTQFGFASSSQPIWLDDVGCTGSEMYLSDC 569

Qy    449 TYDGKAKRTCFRRSDAGVIC---------------------------------------- 468
Db    570 SSSGFGNHNCYHYEDAGVVCQGTSFAIPVRLSNGTRPNEGTVEIFYQNRWGTICDIFWTL 629

Qy    469 ------------------------------------------------------------ 468
Db    630 YDANVVCRQLGYDGAYNATHSSYYGDSGLTTVTAYTDCFGQESQLANCTGFRYPPYIPSW 689

Qy    469 -------------SDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVC 514
Db    690 YCGDNTVAGVMCIGTEDGPAPIRLVGGSEASEGRVEILVQGQWGTICDDIWDQADAEVVC 749

Qy    515 KQLG-CGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIV 573
Db    750 RQLGYLSEGATALQFAHFGEGSGPIVLDDVNCTGLELYVTDCPSSGFYIHNCAHSEDAGV 809

Qy    574 TCSG-----DATWGLRLVGGS---NRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLD 625
Db    810 RCNARGPQTQYTYPIRLVRNGTVITQNEGTVEILHNGTWSAVCDDYWGYTEAVVACHML- 868

Qy    626 CPSSIIGMGLGNASTGY---------GKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSE 676
Db    869 --------GFATAVRAYTRSFHGAVDGDTFLDNVRCAGTEREIFDCYHSSYISRNCDRTQ 920

Qy    677 DVGVICSD--ASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECG 734
Db    921 QAGVACTNYTASEYPIRLVGGSGPHEGRVEIYYQGVWGTVCDDSWGQPDADVVCRQLGYA 980

Qy    735 SAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQ 794
Db    981 NASRATVRAEFGRGTGEIWLDNVACTGFENSLDECRSNGWGDHNCGHREDAGAVCQGELL 1040

Qy    795 P-RLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKG 853
Db   1041 PIRLRDGGSNLEGRVEVYFNKTWGTVCDDFWDLGDATVVCRQLGYPEAVRAEPFARFGSG 1100

Qy    854 NGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVC------SRYTDVRLVNGKS 907
Db   1101 EGPIWLDDVACVGTETSLFECSSSGLGQHNCQHFEDAGVECNDHITPSPVTDLRLAGGTN 1160

Qy    908 QCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYI--GERSVRVWGH 965
Db   1161 ATEGRVEIFFNNTWGTICDDSWDIRDAEVVCRYLGFESAIEALSNGYFGAGDPDQPIWLD 1220

Qy    966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEG 1025

Db   1221 DVDCFGSETIITSCLTSALGEHNCAHYEDAGVRC--YISHPL------------------ 1260

Qy   1026 SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNA 1085
Db   1261 ----------LRLVNGNTEGEGRVEVFHRGSWGTVCDDHWSEIDANIVCKELGFARAISA 1310

Qy   1086 TVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEF----TAL 1141
Db   1311 SGFATFGEGSGTIWLDDVQCDGTEATIFNCPSSGWGNHNCLHSEDASAICTAIPVNPNPV 1370

Qy   1142 RLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGC-GENGVVSLAPLSKT 1200
Db   1371 QLIDGPSVHE--GRLQIYYNNEWGTVCDDQWGYDEANVVCKSLGFPGADPDNSLLNSYGP 1428

Qy   1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDR---IRVRGGDTECSG 1257
Db   1429 GSGHIWLDDVDCRGDEFFIQDCNHADIGENNCGHYEDVGLRCLPNTLDVRLMNGGNTSAG 1488

Qy   1258 RVEIWHAG-SWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCK 1316
Db   1489 RVEVNYNGEEWGTVCDDNWDIKDADVVCRMLDFKSAQSAPREAFFGDGLGVIWFDDFLCS 1548

Qy   1317 GNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLF 1376
Db   1549 GYEDSLLECSHAGVKVHNCRHSEDASVVCS------------------------------ 1578

Qy   1377 ILFLTWCRVQKQKHLPLRVSTRRRG 1401
Db   1579 --MLTVCRDFSQEHGTFQISNRQFG 1601
```

RESULT 8
T30549
hensin - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30549
R;Takito, J.; Hikita, C.; Al-Awqati, Q.
J. Clin. Invest. 98, 2324-2331, 1996
A;Title: Hensin, a new collecting duct protein involved in the in vitro plasticity of
A;Reference number: Z20851; MUID:97096804; PMID:8941650
A;Accession: T30549
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1594 <TAK>
A;Cross-references: EMBL:AF043112; NID:g4105083; PID:g4105084; PIDN:AAD02242.1

```
  Query Match             26.5%;  Score 2156.5;  DB 2;  Length 1594;
  Best Local Similarity   36.6%;  Pred. No. 7.1e-132;
  Matches  475;  Conservative  132;  Mismatches  387;  Indels  303;  Gaps   29;

Qy     46 LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFA---MFRFGQ 102
Db     51 LELRLANGGDRCQGRVEVLYQGSWGTVCDDGWDINDAQVVCRQLGCGMAVSAPGSARFGQ 110

Qy    103 AVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG------------ 149
Db    111 G---PGQIVLDDVSCSGQEPYLWSCHHRGWLSHNCGHQEDAGVICSDAMAMTSPPPDTWP 167

Qy    150 ------------------EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAA 190
Db    168 TTVIYESTPHFPSGLELVFPSGLELRLANGSDRCQGRVEVLYQGSWGTVCDDGWDINDAQ 227

Qy    191 VVCRQLGCPSSFISSGVVNSPAVLR------PIWLDDILCQGNELALWNCRHRGWGNHDC 244
Db    228 VVCRQLGCGMA------VSAPGSARFGQGPGQIVLDDVSCSGQEPYLWSCHHRGWLSHNC 281

Qy    245 SHNEDVTLTCYDS----------------------SDLELRLVGGTNRCMGRVELKIQ 280
Db    282 GHQEDAGVICSDAVPTTTPPPDTWPTTVIYESSPVFPSGLELRLANGSDRCQGRVEVLYQ 341

Qy    281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
```

```
Db    342 GSWGTVCDDGWDINDAQVVCRQLGCGTAVSAPGSARFGQGPGQIVLDDVSCSGQEPYLWS 401

Qy    341 CRHSGTVNFDCLHQNDVSVICSDG-----------------------ADLELRLADGSN 376
Db    402 CHHRGWLSHNCGHQEDAGVICSDAMAMTTPLPDTWPTTVIHESTPVFPSGLELQLANGSD 461

Qy    377 NCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSI 436
Db    462 RCQGRVEVLYQ---GTVCDDGWDINDAQVVCRQLGCGMAVSAPGSARFGQGPGQIVLDDV 518

Qy    437 SCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD-----------------------KA 472
Db    519 SCSGQEPYLWSCHHRGWLSHNCGHQEDAGVICSDAMAMTTPPPDTWPTTVIYESTPHFPS 578

Qy    473 DLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFK 532
Db    579 GLELRLANGSDRCQGRVEVLYQGSWGTVCDDGWDINDAQVVCRQLGCGTAVSAPGSARFG 638

Qy    533 EASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSG--------DATW--- 581
Db    639 QGPGQIVLDDVSCSGQEPYLWSCHHRGWLSHNCGHQEDAGVICSGAMDTTTPLPDTWPTT 698

Qy    582 ------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPS 628
Db    699 VIYESTPVHISGLQLRLVNGSDRCEGRVEVLYQGSWGTVCDDSWDLNDASVVCRQLGCGT 758

Qy    629 SIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDA-SD 687
Db    759 ALSAPASAQFGQSSGSIVLDDVSCSGSEPNLWSCSHRGWLSHNCGHHEDAGVVCSGPDSR 818

Qy    688 MELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTE 747
Db    819 LAVRLVNGSTRCQGRVEVLYRGSWGTVCDDSWDINDASVVCRQLGCGWAVSAPGSARFGQ 878

Qy    748 RTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGR 807
Db    879 GSGSIFLDEVSCSGQEPYLWNCSHRGWLSHNCGHYEDAGVICS---------------- 921

Qy    808 VEVKHADTWRSVC----DSD-FSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKF 862
Db    922 ------DGWTTVTPPAPTTDWWEPTVTTTVGPSSNCGGFLYNATGSFSSP---------- 965

Qy    863 QCEGSETHLALCPIVQHPEDTCIHSREVGV-----VCSRYTDVRLVNGKSQC-------- 909
Db    966 SYPGYYPNNALC-----------VWEIAVPSGYLINLGFSQLRL-EQHSYCNFDYVEIF 1012

Qy    910 DGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHC 969
Db   1013 DGSTDSSLL---GKICN--------------DSGQIFTTSSNR---------------- 1038

Qy    970 LGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALI 1029
Db   1039 ----------MTVL------FRSDISVQNTGFLA------------WYNSFP------ 1062

Qy   1030 CLEDKRLRLVDGDS---RCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNAT 1086
Db   1063 --RDASLRLVSGNSSYGACAGRVEIYHGGRWGTVCDDSWDTQDAQVVCRQLQCGDAVSAP 1120

Qy   1087 VSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC------SEFTA 1140
Db   1121 GGAYFGSGSGPITLDDVNCSGTEATLWQCRSQSWFSHNCGHHEDASVICTGNYGTTTASV 1180

Qy   1141 LRLYSETETESCAGRL--------EVFYNGTWGSVGR 1169
Db   1181 PNISTSNASYSCGGFLSQHSGRFSSPFYPGNYPNNAR 1217
```

RESULT 9
T42721
CRP-ductin-alpha precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 23-Mar-2001
C;Accession: T42721
R;Cheng, H.; Bjerknes, M.; Chen, H.
Anat. Rec. 244, 327-343, 1996
A;Title: CRP-ductin: a gene expressed in intestinal crypts and in pancreatic and hepa
A;Reference number: Z22241; MUID:96362470; PMID:8742698
A;Accession: T42721
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2083 <CHE>
A;Cross-references: EMBL:U37438; NID:g1276646; PID:g1276647; PIDN:AAC52505.1
A;Experimental source: strain BALB/c; jejunal epithelial cells
C;Keywords: transmembrane protein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-2083/Product: CRP-ductin-alpha #status predicted <MAT>

```
  Query Match             24.1%;  Score 1962.5;  DB 2;  Length 2083;
  Best Local Similarity   28.9%;  Pred. No. 3.7e-119;
  Matches  497;  Conservative  164;  Mismatches  503;  Indels  557;  Gaps   41;

Qy     42 NGTD--LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM-- 97
Db     29 DGTDSGLAVRLVNGGDRCQGRVEILYQGSWGTVCDDSWDLNDANVVCRQLGCGLAVSAPG 88

Qy     98 -FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYG------- 149
Db     89 NARFGQG---SGPIVMDDVACGGYEDYLWRCSHRGWLSHNCGHQEDAGVICSDSQTSSPT 145

Qy    150 ----------------------------------EANLGLRLVDGNNSCSGRVEVKFQE 174
Db    146 PGWWNPGGTNNDVFYPTEQTTAEQTTIPDYTPIGTDSGLAVRLVNGGDRCQGRVEILYQG 205

Qy    175 RWGTICDDGWNLNTAAVVCRQLG---------CPSSFISSGVVNSPAVLRPIWLDDILCQ 225
Db    206 SWGTVCDDSWDVSDANVVCRQAGLWLGVSAQEMPDFGQGQGLV----------LDDVACG 255

Qy    226 GNELALWNCRHRGWGNHDCSHNEDVTLTCYDS--------------------------- 257
Db    256 GYENYLWSCSHQGWLSHNCGHQEDAGVICSASQSSSPTPGWWNPGGTNNDVFYPTEQTTA 315

Qy    258 ---SDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGL 314
Db    316 GTDSGLAVRLVNGGDRCQGRVEILYQGSWGTVCDDSWDTNDANVVCRQLGCGWAVSAPGN 375

Qy    315 PHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSD---------- 363
Db    376 AYFGPGSGSIVLDDVACTGHEDYLWRCSHRGWLSHNCGHHEDAGVICSASQSSSPTPDVF 435

Qy    364 ------------------GAD--LELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQ 402
Db    436 YPTDQTTAEQTTVPDYTPIGTDSGLAVRLVNGGDRCQGRVEILYQGSWGTVCDDSWDLND 495

Qy    403 ALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRS 462
Db    496 ANVVCRQLGCGLAVSAPGSARFGQGTGPIVMDDVACGGYEDYLWRCSHRGWLSHNCGHHE 555

Qy    463 DAGVICS-------------------------------DKADLDLRLVGAHSPCYGRLE 490
Db    556 DAGVICSASQSSSPTPDVFYPTDQTTAEQTTVPDYTPIGTDSGLAVRLVNGGDRCQGRVE 615

Qy    491 VKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNES 550
Db    616 ILYQGSWGTVCDDSWDLNDANVVCRQLGCGLAVSAPGSARFGQGTGPIVMDDVACGGYED 675

Qy    551 NIWDCEHSGWGKHNCVHREDVIVTCSG------------------------------D 578
Db    676 YLWRCSHRGWLSHNCGHHEDAGVICSASQSSSPTPDVFYPTDQTTAEQTTVPDYTTIGTE 735

Qy    579 ATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNA 638
Db    736 NSLAVRLENGGDRCQGRVEILYQGSWGTVCDDSWDLNDANVVCRQLGCGLAVSAPGSARF 795

Qy    639 STGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDA------------ 685
```

```
Db     796 GQGTGPIVMDDVACGGYEDYLWRCSHRGWLSHNCGHHEDAGVICSASQSSSPTPDVFYPT 855

Qy     686 ------------------SDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEV 726
                             : : :||  |  || |:||:  ||: | :| : |     | |
Db     856 DQTTVEQTTVPDYTPIGTENSLAVRLENGGDRCQGRVEILYQGSWGTVCDDSWDTKDANV 915

Qy     727 VCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEAS 786
           ||||| || |:        :|    :   |:: :  ||| |  || |     |    |  : :|
Db     916 VCRQLGCGWAVSAPGNAYFGPGSGSIVLDDVACTGHEDYLWSCSHRGWLSHNCGHHEDAG 975

Qy     787 LICS-----AHRQPRL----------------------VGADMPCSG----------- 806
           :|||     :  :| |                      :|   | |
Db     976 VICSDAQIQSTTRPDLWPTTTTPETTTELLTTTPYFDWWTTTSDYSCGGLLTQPSGQFSS 1035

Qy     807 -----------------------RVEVKHADTW----------------------RSVC 820
                                  || |    |                      ||
Db    1036 PYYPSNYPNNARCSWKIVLPNMNRVTVVFTDVQLEGGCNYDYILVYDGPEYNSSLIARVC 1095

Qy     821 DSD-----------------------FSLHAANVLCRELNCGDAISLSVGDHFGKGNG 855
           |                         |  |  : :     :||  :: :    |
Db    1096 DGSNGSFTSTGNFMSVVFITDGSVTRRGFQAHYYSTVSTNYSCGGLLTQPSGQFSSPYYP 1155

Qy     856 LTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQ------- 908
              :     :|         | |              :      | ||    :|||:|  | :
Db    1156 SNYPNNARCSWE----ILVP----------NMNRVTVV---FTDVQLEGGCNYDYILVYD 1198

Qy     909 -----------CDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIG 956
                       |||              ||       |                | :|     :|
Db    1199 GPQYNSSLIARVCDGSN--------GSFTST----------------GNFMSVV---FIT 1231

Qy     957 ERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQP---------- 1006
           : ||      | :: |              ||      ||           |    | ||  |
Db    1232 DGSVTRRGFQAHYYS----------TVSTTPPVPIPTTDDYSCGGLLTLPSGQFSSPHYP 1281

Qy    1007 ------------LFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYH 1053
                       | | :    |:   :       ||       |   |   |   :  ||             |:
Db    1282 SNYPNNARCSWEILVPNMNRVTVAFTDVQLEGG---CNYDYIL-VYDGPE--------YN 1329

Qy    1054 DGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCT--GTESH 1111
                  :||                       |      :|  :|          |           ::: |  :  |   |  ::|
Db    1330 SSLIARVCD-------------------GSNGSFTST-----GNFMSVVFITDGSVTRRGFQAH 1369

Qy    1112 LWQCPSRGW----------GQHDCRH---KEDAGVICS--------------------- 1136
             :     |   :                     ||        |            | ||
Db    1370 YYSTVSTNYSCGGLLTQPSGQFSSPHYPSNYPNNVRCSWEILVPSMNRVTVAFTDVQLEG 1429

Qy    1137 --EFTALRLYSETETES------CAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGE 1188
              |     :  :|      |     |        :||::  |  |    |      :
Db    1430 GCSFDYILVYDGPEYNSSLIAPVCDG-----FNGSFTSTG--NFMS-------------- 1468

Qy    1189 NGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCED---R 1245
             || :     |  |   ||                    |  :     :|:          ||    |
Db    1469 --VVFITDGSVTRRGF-------------------QAYYYSTVSTSLHHPNITGNDSSLA 1507

Qy    1246 IRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGT 1305
           :|:   |     | |||||  :  ||||||||||  :::|  |||:|||||||||:|   :|  ||||:
Db    1508 LRLVNGSNRCEGRVEILYRGSWGTVCDDSWGISDANVVCRQLGCGSALSAPGNAWFGQGS 1567

Qy    1306 GTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCS 1346
           | | |||: | | | || ||:||     |      :| |  |||||  |:
Db    1568 GLIVLDDVSCSGYESHLWNCHHPGWLVHNCRHSEDAGVICA 1608
```

RESULT 10
S56744
mucin (clone pGM7-1) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Oct-1995 #sequence_revision 30-Jan-1998 #text_change 20-Apr-2000
C;Accession: S56744
R;Nunes, D.P.; Keates, A.C.; Afdhal, N.H.; Offner, G.D.
Biochem. J. 310, 41-48, 1995
A;Title: Bovine gall-bladder mucin contains two distinct tandem repeating sequences: evi
A;Reference number: S56744; MUID:95374471; PMID:7646470
A;Accession: S56744
A;Molecule type: mRNA
A;Residues: 1-600 <NUN>
A;Cross-references: GB:S78981; NID:g1042036; PIDN:AAB35069.1; PID:g1042037
C;Superfamily: scavenger receptor cysteine-rich domain homology
F;1-102/Domain: scavenger receptor cysteine-rich domain homology (fragment) <SRC1>
F;126-229/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;253-356/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F;380-483/Domain: scavenger receptor cysteine-rich domain homology <SRC4>

```
  Query Match             13.2%;  Score 1076.5;  DB 2;  Length 600;
  Best Local Similarity   38.5%;  Pred. No. 2.9e-62;
  Matches  223;  Conservative   52;  Mismatches  157;  Indels  147;  Gaps   13;

Qy      48 LRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM---FRFGQAV 104
           ||||||    | | |||  :  | ||||||| |:|   :  |||:|||| :: :    ||||
Db       2 LRLVNGSDRCQGRVEVLYGGSWGTVCDDSWDTNDANVVCRQLGCGWAISAPGDARFGQG- 60

Qy     105 TRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNC----------------- 147
               | | |||| | | |: || | |   | :||| | || | |
Db      61 --SGPIVLDDVGCSGYETYLWSCSHSPWNTHNCGHSEDASVICSASQTQSTVVPDWLYPT 118

Qy     148 --YG-EANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFIS 204
             || |: | ||||:| : | |||||  ::   |||:|||  |: | : ||||||||
Db     119 TDYGTESGLALRLVNGGDRCQGRVEVLYRGSWGTVCDDSWDTNDSNVVCRQLGC------ 172

Qy     205 SGVVNS-PAVLR------PIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTC--- 254
            | || |     |      || |||: | | |  ||:| |  |   |:| |:|| :: |
Db     173 -GWANSAPGNARFGQGSGPIVLDDVGCSGYETYLWSCSHNPWNTHNCGHSEDASVICSAS 231

Qy     255 -----------YDSSD------LELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAAD 297
                       | ::|       | ||||  |::|| ||||:  :| |||||   |:   |:
Db     232 QTQSTVVPDLWYPTTDYGTESGLALRLVNGSDRCQGRVEVLYRGSWGTVCDDSWDTNDAN 291

Qy     298 VVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDV 357
           |||:||||| :   |       ||  :  || | ||| |::|| | |:       :| | |
Db     292 VVCRQLGCGWGISAPGDARFGQGSGPIVLDDVGCSGYETYLWSCSHNPWNTHNCGHSEDA 351

Qy     358 SVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVF 417
           |||||                                      |: ||
Db     352 SVICSA-----------------------------------SQTQSTVV----------- 365

Qy     418 GSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLR 477
                      |:|                          :  :| |      :: | ||
Db     366 ----------PDLW-------------------------YPTTDYGT----ESGLALR 384

Qy     478 LVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGP 537
           ||     | ||:|| |::|| ||||:  | | | | |  |:| |||||||   :    |  |  |||
Db     385 LVNGSDRCQGRVEVLYGGSWGTVCDDSWDTNDANVVCRQLGCGSGISAPGDARFGQGSGP 444

Qy     538 IWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCS 576
            | |||| | | |: :| | || |   ||| | :| | ||
Db     445 ILLDDVGCSGYETYLWSCSHSPWNSHNCGHSKDASVICS 483
```

RESULT 11
A57190
ebnerin precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C;Accession: A57190
R;Li, X.J.; Snyder, S.H.
J. Biol. Chem. 270, 17674-17679, 1995
A;Title: Molecular cloning of Ebnerin, a von Ebner's gland protein associated with ta
A;Reference number: A57190; MUID:95355352; PMID:7629065
A;Accession: A57190
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A;Molecule type: mRNA
A;Residues: 1-1290 <LIA>
A;Cross-references: GB:U32681; NID:g975346; PIDN:AAC52248.1; PID:g975347
C;Superfamily: scavenger receptor cysteine-rich domain homology; Clr/Cls repeat homol
C;Keywords: extracellular protein; transmembrane protein
F;1-77/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC

```
F;142-245/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F;289-392/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;583-689/Domain: C1r/C1s repeat homology <C1R1>
F;713-816/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F;838-944/Domain: C1r/C1s repeat homology <C1R2>

  Query Match             12.9%;  Score 1048.5;  DB 2;  Length 1290;
  Best Local Similarity   29.7%;  Pred. No. 4.5e-60;
  Matches  289;  Conservative  114;  Mismatches  290;  Indels  281;  Gaps   34;

Qy     286 VCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSG 345
             :|   |:    |:|||:|||||  ||    |        |   : ||  |:|  |:|::|| | | |
Db       1 MCDDSWDINDANVVCRQLGCGWALSAPGSAQFGQGLGPIVLDDVACRGHEAYLWSCSHRG 60

Qy     346 TVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALV 405
               :: :| || |  |||||          :  |:   |            ||   :    |   :
Db      61 WLSHNCGHQEDAGVICSD--------SQTSSPTPG---------WW---NPGGTNNDVIY 100

Qy     406 VCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAG 465
             ::          :  :: |:   | |         :    :  |  ||    :       :|:|
Db     101 DTQET--------TETSQTSSPTPDWW-------NHGGTINDVIYD---TQETTEGTDSG 142

Qy     466 VICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHV 525
                     | :|||       |  ||:|:  :|||  |||||  |  |    :|  |||:||||   :
Db     143 --------LAVRLVNGGDRCRGRVEILYQGSWGTVCDDSWDINDANVVCRQLGCGWALSA 194

Qy     526 FGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCS--------- 576
              |    |  :  ||  |  ||||:|  |:|:  :|  |  |  ||     |||  |:||   | |
Db     195 PGSAQFGQGSGSIVLDDVACRGHEAYLWSCSHRGWLSHNCGHQEDAGVICSYSQTSSPTP 254

Qy     577 -----------------GDATWG-------------LRLVGGSNRCSGRLEVYFQGRWG 605
                                |  ::|                 :|||  |  :||  ||:|:  :|| ||
Db     255 DSQTSSPTPGWWNPGGTNNDVSYGPEQTTDATDSGLAVRLVNGGDRCQGRVEILYQGSWG 314

Qy     606 TVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNS 665
             |||||  |::::| |  ||| || |   ::      |    :    |  |  |  ||||:|  |  |:  ||||  :
Db     315 TVCDDSWDTKDANVVCRQLVCGWALSAPGSAHFGQGSGSIVLDDVACTGHEAYLWSCSHR 374

Qy     666 GWGNNDCSHSEDVGVICSDAS-------DM------ELRLVGGSSRCAGKVEVNVQGAVG 712
             || :::|  | ||  ||||||                ||             :::  :     |            :
Db     375 GWLSHNCGHHEDAGVICSDAQTQSTTWPDMWPTTTPETTTDWWTTKYSSSVPTTQFPTI- 433

Qy     713 ILCANGWGMNIAEVVCRQLECGSAIRVS----REPHFTERTLHILMSNSGCTGGEASLWD 768
                |: |         |        || :  :            |::             :|:  |        ||
Db     434 ---ADWWTTPSPEYT-----CGGLLTLPYGQFSSPYYPGS----YPNNARC------LW- 474

Qy     769 CIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSL-- 826
                                          :  |:  :  :|  |:  |                 |: |: |
Db     475 -----------------KIFVSSMNRVTVVFTDVQLEGG-----------CNYDYILVF 505

Qy     827 ----HAANVLCRELNCGDAISLSVGDHFGKG-------------NGLTWAEKFQCE--- 865
                 : :::: |                 |  |  |                        :|          || :
Db     506 DGPENNSSLIAR-----------VCDGFNGSFTSTQNFMSVVFITDGSVTRRGFQADYYS 554

Qy     866 ----GSETHLALCPIV------QHPEDTC--IHSREVGVVCSRYTDVRLVNGKSQCDGQV 913
                  |  |           |||               ||  ||     :   :       |     |   |
Db     555 TPISTSTTSPTTFPIVTDWWTTPSPEYTCGGLLTLPYGQFSSPY--------------- 598

Qy     914 EINVLGHWGSLCDTHWDPEDARVLCR---------------QLSCGT----ALSTTGGK 953
                     :  ||       |  :||  |  :                    ||    |       |      |   :
Db     599 ------YPGSY------PNNARCLWKIFVPSMNRVTVVFTDVQLEGGCNYDYILGFDGPE 646

Qy     954 YIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSV--ICTGSLTQPLF--- 1008
             |             ||       |  |:        :   |                   |  :||    |    ||:|:   |
Db     647 YNSSLIARV------CDGSNGSFTSTQ-----------NFMSVVFITDGSVTRRGFQAD 688

Qy    1009 ----PCLANVSDPYLSAVPEG--SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICD 1062
                    |      :   :   |          :      |     |:|:               ||||:|  :||  |||||  :  |   |        |
Db     689 YYSTPIRTSTTPPTTFPIITGNDSSLV------LRLVNGTNRCEGRVEILYRGSWVPCAD 742

Qy    1063 DGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQ 1122
              | ||::||:|||::||||  |  :|      :|  ||:|||   |  |||::|:|  ||||| |   :||
Db     743 DSWDINDANVVCRQLGCGSALSAPGNAWFGQGSGLIVLDDVSCSGYESHLWNCRHPGWLV 802

Qy    1123 HDCRHKEDAGVICS 1136
              |:||| ||||||||
Db     803 HNCRHVEDAGVICS 816

RESULT 12
T17405
scavenger receptor cysteine-rich protein precursor - sea urchin (Strongylocentrotus p
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C;Accession: T17405
R;Pancer, Z.; Rast, J.P.; Davidson, E.H.
Immunogenetics 49, 773-786, 1999
A;Title: Origins of immunity: transcription factors and homologs of effector genes of
A;Reference number: Z18253; MUID:99328904; PMID:10398804
A;Accession: T17405
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1036 <PAN>
A;Cross-references: EMBL:AF076513; NID:g4164530; PID:g4164531; PIDN:AAD05493.1

  Query Match             10.6%;  Score 866.5;  DB 2;  Length 1036;
  Best Local Similarity   27.1%;  Pred. No. 2.3e-48;
  Matches  229;  Conservative   80;  Mismatches  227;  Indels  309;  Gaps   18;

Qy     312 AGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRL 371
             :|:| |           || ::      :| |                 :|  |            |  |   :|  ||  |
Db     478 SGIPAL--------LDNIAADVVDSIL-------SVLTDA---QQCSTIVQEGLDLALTY 519

Qy     372 ADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVV----------CKQLGCPFSVFGSRR 421
               | :              :: |  :  |     |  |:     ||              |  :  |   |:   ||
Db     520 IPGLD------PYALNLQIYQITGMNGTNQLCQVVADAFQPGDHHCPE-GMTFNECGSGC 572

Qy     422 AKPSNE---ARDI--WINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVIC--SDKADL 474
                 | :       |||               :  |      |     :  |                     |  |      |   |:
Db     573 GPGSCDNLVPRDICPLFCFVGCFCPEGLVKD--------------QDGGDRCIPVDQCQG 618

Qy     475 DLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLG----CGKPMHVFGMTY 530
              :||||    :    ||:|::|  |  |||:|  | |         :|:|||: ||             |  |      |    :
Db     619 AIRLVGGSNEAEGRVEIQYNGVWGTICDDSWGITDASVVCRMLGFQGASGAP----GSAH 674

Qy     531 FKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSN 590
              |  :  :|||  ||||  |  |   |   |:||  |   :|  |||  |  ||    |      |   |:  :|||||||
Db     675 FGQGTGPIQLDDVGCTGAEQTIFDCAHPAFGVHNCAHYEDAGVFCI--ASQDVRLVGGSN 732

Qy     591 RCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDV 650
                 ||:|:  :  | |||:|||  |:      |:|||   |       :         |           |  |  |  ||||
Db     733 EAEGRVEIQYNGVWGTICDDSWDITDASVVCRMLGFQGASGAPGSAQFGQGTGLIQLDDV 792

Qy     651 SCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGA 710
               |  |  |    ::  |  :   :|  ::|:|:|  ||  ||:|   :  |   :||||||:      |:||:     |
Db     793 GCTGAEQTIFDCAHPAFGVHNCAHYEDAGVVCIASQD--VRLVGGSNEAEGRVEIQYNGV 850

Qy     711 VGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCI 770
               | :|  :  ||:    |  |||    |             |   ||  :|  |   :  :  ||||  |  ::::||
Db     851 WGTICDDSWGITDANVVCLMLGFQGASGAPGSAHFGQGTGPIQLDDVGCTGVEQTIFDC- 909

Qy     771 RWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAAN 830
                                                    ||                                      |  :|
Db     910 -------------------AH---------------------------PPFGVH--- 917

Qy     831 VLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREV 890
                                                                                     |  |    :
Db     918 ---------------------------------------------------NCAHYEDA 925

Qy     891 GVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTT 950
             ||||      ||
Db     926 GVVCIVSQDV------------------------------------------------- 935

Qy     951 GGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPC 1010

Db     936 ------------------------------------------------------------ 935
```

```
Qy   1011 LANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDA 1070
                                   |||||  :    ||||| ::| ||||||| ||::||
Db    936 -------------------------RLVDGSNAAEGRVEIQYNGVWGTICDDFWDITDA 969

Qy   1071 HVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKED 1130
          :|||: ||    |   |   || ||:|:||| |||: ||| |    :: |      :| |:| | ||
Db    970 NVVCRMLGFQGASGAPGSAQFGQGTGPIQLDDVGCTGAEQTIFDCAHPPFGVHNCFHSED 1029

Qy   1131 AGVIC 1135
          |||:|
Db   1030 AGVVC 1034
```

RESULT 13
JC5759
brain-specific serine proteinase (EC 3.4.21.-) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 31-Mar-2000
C;Accession: JC5759
R;Yamamura, Y.; Yamashiro, K.; Tsuruoka, N.; Nakazato, H.; Tsujimura, A.; Yamaguchi, N.
Biochem. Biophys. Res. Commun. 239, 386-392, 1997
A;Title: Molecular cloning of a novel brain-specific serine protease with a kringle-like
A;Reference number: JC5759; MUID:98008848; PMID:9344839
A;Accession: JC5759
A;Molecule type: mRNA
A;Residues: 1-761 <YAM>
A;Cross-references: DDBJ:D89871
A;Experimental source: brain
C;Superfamily: brain-specific serine proteinase; scavenger receptor cysteine-rich domain
C;Keywords: glycoprotein; hydrolase; serine proteinase
F;85-157/Domain: kringle-like #status predicted <KRI>
F;163-266/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F;166-266,273-372,386-486/Domain: scavenger receptor cysteine-rich #status predicted <SR
F;513-516/Domain: furin binding #status predicted <FRB>
F;517-755/Domain: trypsin homology <TRY>
F;93,521,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;562,612,711/Active site: His, Asp, Ser #status predicted

```
  Query Match             8.9%;  Score 720.5;  DB 2;  Length 761;
  Best Local Similarity   31.8%;  Pred. No. 4.8e-39;
  Matches  191;  Conservative   80;  Mismatches  233;  Indels   97;  Gaps   19;

Qy    451 DGKAKRTCFRRSDAGVI------CSDKADLD-LRLVGAHSPCYGRLEVKYQGEWGTVCHD 503
          ||  :  || |: |  :        |      |  :|||| :|    ||:|:  :.|:|||:| |
Db    134 DGSGRPWCFYRNAQGKVDWGYCDCGQGPALPVIRLVGGNSGHEGRVELYHAGQWGTICDD 193

Qy    504 RWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKH 563
          :|    :| |:|:|||          :  :| | |||| ||:| | ||| :| , |  | ||:|
Db    194 QWDNADADVICRQLGLSGIAKAWHQAHFGEGSGPILLDEVRCTGNEMSIEQCPKSSWGEH 253

Qy    564 NCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQ 623
          || |:||  |:|        :|| || :   ||||||::|:||||||||       |·|
Db    254 NCGHKEDAGVSCVPLTDGVIRLAGGKSTHEGRLEVYYKGQWGTVCDDGWTEMNTYVAC-- 311

Qy    624 LDCPSSIIGMGLGNAST-----GYGK-IWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
            ::|   |  |:      |  : |||||||| | |          ||  :|||| ||
Db    312 -----RLLGFKYGKQSSVNHFDGSNRPIWLDDVSCSGKEVSFIQCSRRQWGRHDCSHRED 366

Qy    678 VGVICSDASD-------MELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQ 730
          ||: |   ||         :||| | ::   |:||| | |   |  |:| :||    | |:|||
Db    367 VGLTCYPDSDGHRLSPGFPIRLVDGENKKEGRVEVFVNGQWGTICDDGWTDKHAAVICRQ 426

Qy    731 LECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICS 790
          |      |        :| |    | | |   ||| | :| ||:: :   :   |    : :| :||
Db    427 LGYKGPARARTMAYFGEGKGPIHMDNVKCTGNEKALADCVKQDIGRHNCRHSEDAGVICD 486

Qy    791 ------------------------AHRQPRLVGADMPCSG---------RVEVKHADTWRS 818
                                  || ||::|  :      |          |:    |   |   |
Db    487 YLEKKASSSGNKEMLSSGCGLRLLHRRQKRIIGGNNSLRGAWPWQASLRLESAHGD-GRL 545

Qy    819 VC-----DSDFSLHAANVLCRELNCGDAISLSVGDH-------FGKGNG---LTWAEKFQ 863
          :|        | : | ||:    |   |    : :: |||:          | :  |    :      ::
Db    546 LCGATLLSSCWVLTAAHCFKRYGNNSRSYAVRVGDYHTLVPEEFEQEIGVQQIVIHRNYR 605

Qy    864 CEGSETHLALCPIVQHPEDTCIH-SREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWG 922
           : |:   :||   : | |  : |   |   |   |      |       |  ||      :: | ||
Db    606 PDRSDYDIALVRL-QGPGEQCARLSTHVLPACLPLWRERPQKTASNC------HITG-WG 657

Qy    923 SLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN---ESLLDNC 979
                  :    |    | |    |:        ::   ||    ::    |   | ||    :: :|:|
Db    658 ---------DTGRAYSRTLQ-QAAVPLLPKRFCKERYKGLFTGRMLCAGNLQEDNRVDSC 707

Qy    980 Q 980
          |
Db    708 Q 708
```

RESULT 14
JC4361
scavenger receptor Cys-rich epidermal growth factor precursor - sea lamprey
C;Species: Petromyzon marinus (sea lamprey)
C;Date: 08-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 21-Jul-2000
C;Accession: JC4361
R;Mayer, W.E.; Tichy, H.
Gene 164, 267-271, 1995
A;Title: A cDNA clone from the sea lamprey Petromyzon marinus coding for a scavenger
A;Reference number: JC4361; MUID:96069593; PMID:7590341
A;Accession: JC4361
A;Molecule type: mRNA
A;Residues: 1-918 <MAY>
A;Cross-references: GB:U20652; NID:g790233; PIDN:AAA90990.1; PID:g790234
C;Comment: This protein is rich in cysteine and plays a role in intercellular contact
C;Superfamily: unassigned EGF-related proteins; EGF homology; scavenger receptor cyst
C;Keywords: glycoprotein; growth factor; receptor; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-918/Product: scavenger receptor Cys-rich epidermal growth factor #status predict
F;55-157/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F;179-212/Domain: EGF homology <EG1>
F;234-267/Domain: EGF homology <EG2>
F;295-328/Domain: EGF homology <EG3>
F;356-389/Domain: EGF homology <EG4>
F;410-443/Domain: EGF homology <EG5>
F;463-565/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;876-895/Domain: transmembrane #status predicted <TMM>
F;896-918/Domain: intracellular #status predicted <INT>
F;40,545,575,585,814/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
  Query Match             8.2%;  Score 665;  DB 2;  Length 918;
  Best Local Similarity   25.7%;  Pred. No. 2.4e-35;
  Matches  256;  Conservative   87;  Mismatches  336;  Indels  318;  Gaps   41;

Qy    471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG----KPMHV- 525
          :||| :|||   |  | ||:|:  |  |||||  |  | :| |||:|||||      | |
Db     53 EADLPVRLVQGGS-CSGRVEILYGGRWGTVCDDNWGIEDATVVCRQLGCGYAVYAPTHAR 111

Qy    526 FGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRL 585
          |||     |||||:|:| | |:|| :   |  :||   |:| | ||  | || ||
Db    112 FGM-----GSGPIWMDEVYCSGSESYLSHCLFNGWNSHDCSHGEDASVICSHDA------ 160

Qy    586 VGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTG-YGK 644
                    | :|| | |                                  :|| ||
Db    161 ---------------GSYGTTVDAG--------------------------PTTGYYGA 178

Qy    645 IWLDDVSCDGDESDL-------WSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSS 697
             :         |    ::|                 ||:|   |  :|  :                 | |     |     :
Db    179 CYSHPCQNGGSCTNLPYYPYYRCHCRSS-WTGTNCEYYYHESTTLPGAEDTTHSSVCQRN 237

Qy    698 FCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGS----------AIRVSREPHFT 746
          |       |     |        :    |      |:                  |    :   |
Db    238 FCQH----------GGTCVEYWNYPYYQCYCPSYWWGTNCEYYGAHSTDAPESTTLPRAE 287

Qy    747 ERTLHILMSNSGCTGGEASLWDCIR-WEWKQTACH-------LNMEASLICSAHRQPRLV 798
           : |    :   : |   |     |:  |  :    |:            | |          ||
Db    288 DTTHSSVCQRNPCQHGGT----CVEYWNYPYYQCYCPSYWWGTNCE---YYGAH------ 334
```

```
Qy    799 GADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTW 858
              | | |  :         |  |||
Db    335 STDAPESTTLPGAEDTTHSSVC-------------------------------------- 356

Qy    859 AEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVL 918
              | |        || |:: |    |          |  |
Db    357 ----QXEPMPAWRNLCGILELPYYQC--------YCPSY-------------------- 383

Qy    919 GHWGSLCDTHW-DPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLD 977
           ||: |: :     ||  |            ||| | |              |  :
Db    384 -WWGTNCEYYGAHSTDAPEL-----------TTG--YYGA------------CYSH----- 413

Qy    978 NCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPC---------------LANVSDPYLSAV 1022
                      || :| :     ||      | : |                  : | | | :
Db    414 ---------PCQNGGS----CTNLPYPYPYRCHCRSSWTGTNCEYYYYGSTSSPELDST 459

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
                  |  :||| |    | ||||:|: | |||:||| ||  || |||::| || |
Db    460 -------LSPDVPVRLV-GGGWCQGRVEVYYAGSWGTVCDDSWDRQDAEVVCRQLSCGYA 511

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
            :|  :||||  |||  |||||:||  |  || |  :|    |||:|:| | ||| ||||     |
Db    512 VSAPSNAHFGVGSGQIWLDDVNCYGQESSLARCNHSGWGRHNCGHNEDASVICS-----R 566

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
            | | :                 |::   : | |  :     |    |  | :  | |||:  |
Db    567 APSTTPS---------------GNI--TSPTNAPSNSTSRIFHCNSNYMELLIPLSEISS 609

Qy   1203 GFMWVDDI--QCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260
            : | ||  | |  | : :   :      |:  |          :    |  || ||
Db    610 RGLSVYDIRLQDPSCHATAYGQYA-----RLYGP-----------LNGCGSRTETSGDTI 653

Qy   1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALR-----DASFGQGTGTIWLDDMRC 1315
           |:  ::| |  :           : |      | |:|      : | |        | |:|
Db    654 IYMNTAYGYVSPNGAKRLRVPMFCYMGSSGRAMATFTPRVQDEVSMG-------LFDLRV 706

Qy   1316 K--GNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALIL-----SSIF 1368
            :    |: |    :: |:    : |     | |          :| :::  | | |      |
Db    707 RFYMNQFFHQPIYSYPY-DVELGEPLYAEV--------ALTSAARDLQLFLETCKVSPYP 757

Qy   1369 GLLLLVLFILFLTWCRVQ-KQKHLPLRVSTRRRGSLE 1404
            |        | |    .|||          |: ||| |  :
Db    758 GAPDSASFFLLQNGCRVDPSYSDYNSGVNDRRRFSFD 794
```

RESULT 15
I49100
mscd6 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49100
R;Whitney, G.; Bowen, M.; Neubauer, M.; Aruffo, A.
Mol. Immunol. 32, 89-92, 1995
A;Title: Cloning and characterization of murine CD6.
A;Reference number: I49100; MUID:95174761; PMID:7870060
A;Accession: I49100
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-626 <RES>
A;Cross-references: EMBL:U12434; NID:g761719; PIDN:AAA64867.1; PID:g761720
C;Genetics:
A;Gene: Cd6
C;Superfamily: scavenger receptor cysteine-rich domain homology
F;42-155/Domain: scavenger receptor cysteine-rich domain homology <SRC>
F;157-259/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F;261-360/Domain: scavenger receptor cysteine-rich domain homology <SRC3>

```
  Query Match             7.7%;  Score 626.5;  DB 2;  Length 626;
  Best Local Similarity   28.6%;  Pred. No. 4.9e-33;
  Matches  164;  Conservative   60;  Mismatches  157;  Indels  193;  Gaps   14;

Qy    900 VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERS 959
           :||||| |:| | |:: :|  |  :|  ||:        :|: |:||
Db     45 IRLVNGSSRCSGSVKV-LLESWEPVCAAHWNRAATEAVCKALNCG-------------- 88

Qy    960 VRVWGHRFHCLGNESLLDNCQMTVLGAPP------CIHGNTVSVICTGSLTQPLFPC--- 1010
                                 |: ::| | · |          ||| |    |    |    |
Db     89 ----------------DSGKVTYLMPPTSELPPGATSGNTSSAGNTTWARAPTERCRGA 131

Qy   1011 ---LANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDL 1067
               | |    |:          : | |:: :||||| ||||||||:     | |||:||| |||
Db    132 NWQFCKVQDQECSSDRRLVWVTCAENQAVRLVDGSSRCAGRVEMLEHGEWGTVCDDTWDL 191

Qy   1068 SDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRH 1127
            |||||::| || |   |        ||   | |||   | :||:|||::|| || |    |: | |
Db    192 QDAHVVCKQLKCGWAVKALAGLHFTPGQGPIHRDQVNCSGTEAYLWDCPGRPGDQY-CGH 250

Qy   1128 KEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCG 1187
            ||||||:|||
Db    251 KEDAGVVCSE-------------------------------------------------- 260

Qy   1188 ENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIR 1247
                                                               ::|       |
Db    261 -------------------------------------------------HQSW-------R 265

Qy   1248 VRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGT 1307
            : ||   | |:||::  | | ||||   |   :||:|:|: ||||||:| |
Db    266 LTGGIDSCEGQVEVYFRGVWSTVCDSEWYPSEAKVLCRSLGCGSAVARPRGLPHS----- 320

Qy   1308 IWLDD---MRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASS------- 1357
              ||      ||| |  |  |  |  :    : |        | | ||| | : || |:
Db    321 --LDGRMYYSCKGQEPALSTCSWRFNNSNLCSQSRAARVVCSG-SQRHLNLSTSEVPSRV 377

Qy   1358 -----------------GHLALILSSIFGLLLLV--LFILFLTWCRVQKQKHLPLRVST 1397
                                 |   |||   : |:||||   :||: |     | : |    ||   |:
Db    378 PVTIESSVPVSVKDKDSQGLTLLILCIVLGILLLVSTIFIVILL-LRAKGQYALPASVNH 436

Qy   1398 RRRGSLEE---NLFHEMETCLKREDPHGTRTSDD 1428
            :: :  :    | :| :   : :||   | :::
Db    437 QQLSTANQAGINNYHPVPITIAKEDSQRHRVTEE 470
```

Search completed: May 12, 2003, 06:43:42
Job time : 48 secs

```
OM protein - protein search, using sw model

Run on:         May 12, 2003, 05:30:12 ; Search time 21 Seconds
                                          (without alignments)
                                          2869.767 Million cell updates/sec

Title:          US-09-759-130B-381
Perfect score:  8138
Sequence:       1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:      112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_40:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3082	37.9	1436	1	WC11_BOVIN	P30205 bos taurus
2	822.5	10.1	875	1	NETR_HUMAN	P56730 homo sapien
3	719.5	8.8	761	1	NETR_MOUSE	O08762 mus musculu
4	691.5	8.5	347	1	CD5L_HUMAN	O43866 homo sapien
5	664.5	8.2	352	1	CD5L_MOUSE	Q9qwk4 mus musculu
6	633.5	7.8	757	1	LOL4_MOUSE	Q924c6 mus musculu
7	629	7.7	754	1	LOL3_MOUSE	Q9z175 mus musculu
8	622.5	7.6	753	1	LOL3_HUMAN	P58215 homo sapien
9	618	7.6	756	1	LOL4_HUMAN	Q96jb6 homo sapien
10	614.5	7.6	665	1	CD6_MOUSE	Q61003 mus musculu
11	607.5	7.5	774	1	LOL2_HUMAN	Q9y4k0 homo sapien
12	592	7.3	668	1	CD6_HUMAN	P30203 homo sapien
13	520	6.4	532	1	SPER_STRPU	P16264 strongyloce
14	298	3.7	483	1	MRCO_MESAU	Q9wub9 mesocricetu
15	290.5	3.6	518	1	MTCO_MOUSE	Q60754 mus musculu
16	290	3.6	458	1	MSRE_MOUSE	P30204 mus musculu
17	288.5	3.5	520	1	MRCO_HUMAN	Q9uew3 homo sapien
18	288	3.5	451	1	MSRE_HUMAN	P21757 homo sapien
19	279.5	3.4	453	1	MSRE_BOVIN	P21758 bos taurus
20	279.5	3.4	454	1	MSRE_RABIT	Q05585 oryctolagus
21	264	3.2	494	1	CD5_MOUSE	P13379 mus musculu
22	258.5	3.2	2470	1	NTC2_MOUSE	O35516 mus musculu
23	258	3.2	5376	1	ZAN_MOUSE	O88799 mus musculu
24	254.5	3.1	495	1	CD5_HUMAN	P06127 homo sapien
25	248.5	3.1	2471	1	NTC2_HUMAN	Q04721 homo sapien
26	247.5	3.0	2471	1	NTC2_RAT	Q9qw30 rattus norv
27	246.5	3.0	2531	1	NTC1_MOUSE	Q01705 mus musculu
28	244.5	3.0	2871	1	FBN1_MOUSE	Q61554 mus musculu
29	242.5	3.0	491	1	CD5_RAT	P51882 rattus norv
30	240.5	3.0	1238	1	JAG2_HUMAN	Q9y219 homo sapien
31	239.5	2.9	4544	1	LRP1_HUMAN	Q07954 homo sapien
32	239	2.9	2531	1	NTC1_RAT	Q07008 rattus norv
33	233	2.9	1700	1	BAR3_CHITE	Q03376 chironomus
34	231.5	2.8	2871	1	FBN1_HUMAN	P35555 homo sapien
35	230	2.8	2556	1	NTC1_HUMAN	P46531 homo sapien
36	230	2.8	2703	1	NOTC_DROME	P07207 drosophila
37	230	2.8	2871	1	FBN1_BOVIN	P98133 bos taurus
38	229.5	2.8	1202	1	JAG2_RAT	P97607 rattus norv
39	224.5	2.8	2871	1	FBN1_PIG	Q9tv36 sus scrofa
40	224	2.8	2907	1	FBN2_MOUSE	Q61555 mus musculu
41	223.5	2.7	495	1	CD5_BOVIN	P19238 bos taurus
42	223.5	2.7	2911	1	FBN2_HUMAN	P35556 homo sapien
43	220	2.7	4543	1	LRP1_CHICK	P98157 gallus gall
44	216.5	2.7	2318	1	NTC3_MOUSE	Q61982 mus musculu
45	216	2.7	1247	1	JAG2_MOUSE	Q9qye5 mus musculu

ALIGNMENTS

```
RESULT 1
WC11_BOVIN
ID   WC11_BOVIN     STANDARD;      PRT;  1436 AA.
AC   P30205;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Antigen WC1.1 precursor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=93056489; PubMed=1431105;
RA   Wijngaard P.L.J., Metzelaar M.J., Machugh N.D., Morrison W.I.,
RA   Clevers H.C.;
RT   "Molecular characterization of the WC1 antigen expressed specifically
RT   on bovine CD4-CD8-gamma delta T lymphocytes.";
RL   J. Immunol. 149:3273-3277(1992).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: CONTAINS 11 SRCR DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X63723; CAA45255.1; -.
DR   PIR; S19913; S19913.
DR   PIR; A46496; A46496.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 11.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 11.
DR   PROSITE; PS00420; SRCR_1; 4.
DR   PROSITE; PS50287; SRCR_2; 11.
KW   Antigen; Repeat; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26   1436       ANTIGEN WC1.1.
FT   DOMAIN       28    131       SRCR 1.
FT   DOMAIN      134    234       SRCR 2.
FT   DOMAIN      239    340       SRCR 3.
FT   DOMAIN      376    476       SRCR 4.
FT   DOMAIN      481    581       SRCR 5.
FT   DOMAIN      586    686       SRCR 6.
FT   DOMAIN      689    789       SRCR 7.
FT   DOMAIN      794    895       SRCR 8.
FT   DOMAIN      931   1031       SRCR 9.
FT   DOMAIN     1036   1136       SRCR 10.
FT   DOMAIN     1155   1255       SRCR 11.
```

```
FT   CARBOHYD    162    162       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    424    424       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    444    444       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    499    499       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    531    531       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    717    717       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    799    799       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    897    897       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    979    979       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    999    999       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1054   1054      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1086   1086      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1173   1173      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1214   1214      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1393   1393      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1436 AA;  154196 MW;  6C72E5FDC6BA088C CRC64;

  Query Match             37.9%;  Score 3082;  DB 1;  Length 1436;
  Best Local Similarity   41.7%;  Pred. No. 1.2e-208;
  Matches  613;  Conservative  208;  Mismatches  495;  Indels  154;  Gaps   23;

Qy      28 CILLLNSCFLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCK 87
           |:|||      :    |  |||||  :|   | |  |||| ||:||||     |      ::|||:
Db      13 CVLLLG-----TMVGGQALELRLKDGVHRCEGRVEVKHQGEWGTVDGYRWTLKDASVVCR 67

Qy      88 QLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQH---REWGSHNCYHGE 141
           ||||  :          ||  :    | |||   || | ||  : :|:|    ::: :      ||
Db      68 QLGCGAAIGFPGGAYFGPGL---GPIWLLYTSCEGTESTVSDCEHSNIKDYRNDGYNHGR 124

Qy     142 DVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS 201
           | || | |       :||  |:  ||||||    | |   :  |   :  | ||  ::| :|||   :
Db     125 DAGVVCSG----FVRLAGGDGPCSGRVEVHSGEAWIPVSDGNFTLATAQIICAELGCGKA 180

Qy     202 FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE 261
                |          :|  ::   |:|  |  ||  |         | |:       :  |    |::
Db     181 VSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPCPGGTCHHSGSAQVVCSAYSEVR 240

Qy     262 LRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGS 321
           | :   |:::| |:||: | |:|   :|    |: | |:|:|:|||| |:     | |||
Db     241 L-MTNGSSQCEGQVEMNISGQWRALCASHWSLANANVICRQLGCGVAISTPGGPHLVEEG 299

Qy     322 DVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGR 381
           | :         ||| |||||  |     :          ||  |  |  |||||
Db     300 DQILTARFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGN----------------- 342

Qy     382 VEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGN 441
            ::::  |                     |  ||       ::|:  |
Db     343 -QIQVLPQ--------------------CNDSV-----SQPTGSAAS----------- 363

Qy     442 ESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVC 501
                           |:   |||         ||||      || ||:|:   || |||:|
Db     364 -------------------EDSAPYCSDSR--QLRLVDGGGPCAGRVEILDQGSWGTIC 401

Qy     502 HDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWG 561
            |  |    :|  |||:|||||: ::   |    :|    |||||||:::| | ||::| |    |||
Db     402 DDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKESHVWRCPSRGWG 461

Qy     562 KHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVC 621
           :||| |::| | ||          ||:|       :|:| |||:: | ||:|| :          : :|
Db     462 QHNCRHKQDAGVICS--EFLALRMVSEDQQCAGWLEVFYNGTWGSVCRNPMEDITVSTIC 519

Qy     622 SQLDCPSSIIGMGLGNAST----GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
            || |     |  |:|        |:   |:| |   :: || |  :  |  | ||  |:
Db     520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWNYNSCSPKEE 575

Qy     678 VGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAI 737
            : |:|:      ::|||  | ||:|:||: ||:  | :|   |  :  :  ||||:|| |:| |:
Db     576 AYIWCADS--RQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVCKQLGCGEAL 633

Qy     738 RVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRL 797
             :    |   :  | :    | | |:  :| |  | |  |:|  | :     :| :|||     ||
Db     634 DATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAGVICSGF--VRL 691

Qy     798 VGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLT 857
            | | ||||||||    :  |  | | :|:|  | |:| || || || |:|:        | : :|
Db     692 AGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESDGQV 751

Qy     858 WAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLV-NGKSQCDGQVEIN 916
           |||:|:|:|  |    | ||   |   ||:||  |||| ||:|:|:|: || |||:||||:
Db     752 WAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMKNGTSQCEGQVEMK 811

Qy     917 VLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLL 976
            : |  | :||  :||    :|  |:||||  ||  |:|||  | :: |    ::   :|||  ||| |
Db     812 ISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQISTAQFHCSGAESFL 871

Qy     977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036
            :| :| || | | |||| ||||:|: ||  |   |    :| |    ||   | |:  | : ::|
Db     872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSAASEESSPYCSDSRQL 931

Qy    1037 RLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSG 1096
           |||||    | |||||    | ||||||| ||| || |||::|||| | ||| ||||| |||
Db     932 RLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQLGCGEALNATGSAHFGAGSG 991

Qy    1097 PIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRL 1156
           |||||||||||| |||:|:||||||||:|||||||||||||||||||||| |||:   |   | :  ||| |
Db     992 PIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVICSEFLALRMVS--EDQQCAGWL 1049

Qy    1157 EVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTH 1216
           |||||||||| |    :        ::||||||||::|  :: :    ||    ||| ||| |
Db    1050 EVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTSVGLREGSRPRWVDLIQCRKMD 1109

Qy    1217 ISIWQCLSAPWERRISSPAEETWITCE---------------DRIRVRGGDTECSGRVE 1260
            |:||| | ||:     || || :|:||                    :::|:|||:|:|||||||
Db    1110 TSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTAACTDREKLRLRGGDSECSGRVE 1169

Qy    1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
           :|| ||||||||||| |||||||||||| || |:| |:|| | |:||||:::| | ||
Db    1170 VWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEAVRSAAFGPGNGSIWLDEVQCGGRES 1229

Qy    1321 FLWDCHAKPWGQSDCGHKEDAGVRCSG---------------QSLKSLNASSGHLALIL 1364
            |||| |:||||||| |:|||||||||                    || :  :  | | |||
Db    1230 SLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLPTTTAGTRTTSNSLPGIFSLPGVLCLIL 1289

Qy    1365 SSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTR 1424
            |:  |:|::|    | | | ::       |   :    :| |.:: |::  | ::: |:
Db    1290 GSLLFLVLVILVTQLLRW-RAER------RALSSYEDALAEAVYEELDYLLTQKEGLGSP 1342

Qy    1425 TS-DDTPNHGCEDASDTSLLGVLPASEATK 1453
            | |:  :|| :  : |     |:   :
Db    1343 DQMTDVPDENYDDAEEVPVPGTPSPSQGNE 1372


RESULT 2
NETR_HUMAN
ID   NETR_HUMAN     STANDARD;      PRT;   875 AA.
AC   P56730; Q9UP16;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Neurotrypsin precursor (EC 3.4.21.-) (Motopsin) (Leydin).
GN   PRSS12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98201705; PubMed=9540828;
RA   Proba K., Gschwend T.P., Sonderegger P.;
RT   "Cloning and sequencing of the cDNA encoding human neurotrypsin.";
RL   Biochim. Biophys. Acta 1396:143-147(1998).
RN   [2]
```

```
RP   SEQUENCE OF 615-875 FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=99203523; PubMed=10103056;
RA   Poorafshar M., Hellman L.;
RT   "Cloning and structural analysis of leydin, a novel human serine
RT   protease expressed by the Leydig cells of the testis.";
RL   Eur. J. Biochem. 261:244-250(1999).
CC   -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
CC       ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
CC       LEARNING AND MEMORY OPERATIONS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Brian and Leydig cells of the testis.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ001531; CAA04816.1; -.
DR   EMBL; AF077298; AAD25919.1; -.
DR   HSSP; P00763; 1DPO.
DR   Genew; HGNC:9477; PRSS12.
DR   MIM; 606709; -.
DR   MEROPS; S01.237; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00530; SRCR; 4.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00202; SR; 4.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS00420; SRCR_1; 3.
DR   PROSITE; PS50287; SRCR_2; 4.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    875       NEUROTRYPSIN.
FT   DOMAIN       23     92       PRO-RICH.
FT   DOMAIN       93    165       KRINGLE.
FT   DOMAIN      170    271       SRCR 1.
FT   DOMAIN      280    381       SRCR 2.
FT   DOMAIN      387    487       SRCR 3.
FT   DOMAIN      500    601       SRCR 4.
FT   DOMAIN      619    875       SERINE PROTEASE.
FT   DOMAIN      619    630       ZYMOGEN ACTIVATION REGION.
FT   ACT_SITE    630    631       REACTIVE BOND (POTENTIAL).
FT   ACT_SITE    676    676       CHARGE RELAY SYSTEM.
FT   ACT_SITE    726    726       CHARGE RELAY SYSTEM.
FT   ACT_SITE    825    825       CHARGE RELAY SYSTEM.
FT   DISULFID    619    750       POTENTIAL.
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    683    683       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    663    663       A -> V (IN REF. 2).
FT   CONFLICT    701    701       E -> V (IN REF. 2).
FT   CONFLICT    839    841       VVY -> AAL (IN REF. 2).
SQ   SEQUENCE   875 AA;  97011 MW;  B66EC946DC208DC8 CRC64;
```

```
  Query Match             10.1%;  Score 822.5;  DB 1;  Length 875;
  Best Local Similarity   29.6%;  Pred. No. 3.3e-50;
  Matches  229;  Conservative   96;  Mismatches  307;  Indels  141;  Gaps   26;

Qy        239 WGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAADV 298
              || || |                      :|| || |   | ||:   | |||||   |::: | |
Db        160 WGYCDCRHG-------------SVRLRGGKNEFEGTVEVYASGVWGTVCSSHWDDSDASV 206

Qy        299 VCKQL-----GCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLH 353
              :| ||      |       |:||     |   ::    | | |:|   :   |          |
Db        207 ICHQLQLGGKGIAKQTPFSGL-----GLIPIYWSNVRCRGDEENILLCEKDIWQGGVCPQ 261

Qy        354 QNDVSVIC--SDGADLE-LRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQL 410
              :   :| |  | |       :||| ||:    ||||:     || |:||  | :   | |:|:||
Db        262 KMAAAVTCSFSHGPTFPIIRLAGGSSVHEGRVELYHAGQWGTVCDDQWDDADAEVICRQL 321

Qy        411 GCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470
              |         :|      :  : :: : ||||| ::   |         :   |   : |||| |:
Db        322 GLSGIAKAWHQAYFGEGSGPVMLDEVRCTGNELSIEQCPKSSWGEHNCGHKEDAGVSCTP 381

Qy        471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLG--CGKPMHVFGM 528
                |  :|| |         ||||| |:|:||||| | |:    |   |||:|||   ||
Db        382 LTDGVIRLAGGKGSHEGRLEVYYRGQWGTVCDDGWTELNTYVVCRQLGFKYGKQA---SA 438

Qy        529 TYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTC--SGDA-----TW 581
               :|:|::||||||||| | |:     |      ||:|:| ||||| : |     |:       :
Db        439 NHFEESTGPIWLDDVSCSGKETRFLQCSRRQWGRHDCSHREDVSIACYPGGEGHRLSLGF 498

Qy        582 GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTG 641
               :||: | |:  ||:||:   |:|||:|||||   | |||:| ||                :       |
Db        499 PVRLMDGENKKEGRVEVFINGQWGTICDDGWTDKDAAVICRQLGYKGPARARTMAYFGEG 558

Qy        642 YGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAG 701
                | | :|:| | |:|  |   |      | ::| |||| ||||      |    :    |:|
Db        559 KGPIHVDNVKCTGNERSLADCIKQDIGRHNCRHSEDAGVIC----DYFGKKASGNS---- 610

Qy        702 KVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTG 761
                   | :     :                         || :        |   |:   :     |
Db        611 ----NKESLSSV------------------CGLRL-------LHRRQKRIIGGKNSLRG 640

Qy        762 GEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCD 821
              |         | | | | :   |       |:|    ||:       |                ::
Db        641 G---------WPW-QVSLRLK-------SSHGDGRLL-----CGA-----------TLLS 667

Qy        822 SDFSLHAANVLCRELNCGDAISLSVGDH-------FGKGNG---LTWAEKFQCEGSETHL 871
              | : | ||:    |   |     : :: |||:         | :   |      ::: : |:   :
Db        668 SCWVLTAAHCFKRYGNSTRSYAVRVGDYHTLVPEEFEEEIGVQQIVIHREYRPDRSDYDI 727

Qy        872 ALCPIVQHPEDTCIH-SREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWD 930
              ||   :  | ||: |     |   |   |           |             |      :   ||
Db        728 ALVRL-QGPEEQCARFSSHVLPACLPLWRERPQKTASNC-------YITGWG-------- 771

Qy        931 PEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN---ESLLDNCQ 980
               :  |   | |   |:       ::  ||      :  |  |  ||        :|:||
Db        772 -DTGRAYSRTLQ-QAAIPLLPKRFCEERYKGRFTGRMLCAGNLHEHKRVDSCQ 822
```

```
RESULT 3
NETR_MOUSE
ID   NETR_MOUSE      STANDARD;      PRT;   761 AA.
AC   O08762;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Neurotrypsin precursor (EC 3.4.21.-) (Motopsin) (Brain-specific serine
DE   protease 3) (BSSP-3).
GN   PRSS12 OR BSSP3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Brain;
RX   MEDLINE=97401523; PubMed=9245503;
RA   Gschwend T.P., Krueger S.R., Kozlov S.V., Wolfer D.P., Sonderegger P.;
RT   "Neurotrypsin, a novel multidomain serine protease expressed in the
RT   nervous system.";
RL   Mol. Cell. Neurosci. 9:207-219(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98008848; PubMed=9344839;
RA   Yamamura Y., Yamashiro K., Tsuruoka N., Nakazato H., Tsujimura A.,
RA   Yamaguchi N.;
RT   "Molecular cloning of a novel brain-specific serine protease with a
RT   kringle-like structure and three scavenger receptor cysteine-rich
RT   motifs.";
RL   Biochem. Biophys. Res. Commun. 239:386-392(1997).
CC   -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
CC       ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
CC       LEARNING AND MEMORY OPERATIONS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: MOST ABUNDANT IN CEREBRAL CORTEX, HIPPOCAMPUS
CC       AND AMYGDALA.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y13192; CAA73646.1; -.
DR   EMBL; D89871; BAA23986.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.237; -.
DR   MGD; MGI:1100881; Prss12.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00530; SRCR; 3.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00202; SR; 3.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS00420; SRCR_1; 3.
DR   PROSITE; PS50287; SRCR_2; 3.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    761       NEUROTRYPSIN.
FT   DOMAIN       85    157       KRINGLE.
FT   DOMAIN      166    267       SRCR 1.
FT   DOMAIN      273    373       SRCR 2.
FT   DOMAIN      386    487       SRCR 3.
FT   DOMAIN      505    761       SERINE PROTEASE.
FT   DOMAIN      505    516       ZYMOGEN ACTIVATION REGION.
FT   ACT_SITE    516    517       REACTIVE BOND (POTENTIAL).
FT   ACT_SITE    562    562       CHARGE RELAY SYSTEM.
FT   ACT_SITE    612    612       CHARGE RELAY SYSTEM.
FT   ACT_SITE    711    711       CHARGE RELAY SYSTEM.
FT   DISULFID    505    636       POTENTIAL.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    521    521       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    569    569       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   761 AA;  84118 MW;  DF507B03712164E6 CRC64;

  Query Match             8.8%;  Score 719.5;  DB 1;  Length 761;
  Best Local Similarity   31.8%;  Pred. No. 4.8e-43;
  Matches  191;  Conservative   80;  Mismatches  233;  Indels   97;  Gaps   19;

Qy        451 DGKAKRTCFRRSDAGVI------CSDKADLD-LRLVGAHSPCYGRLEVKYQGEWGTVCHD 503
              ||  :  || |:  | :      |     |  :|||| :|   ||:|: : |:|||:| |
Db        134 DGSGRPWCFYRNAQGKVDWGYCDCGQGPALPVIRLVGGNSGHEGRVELYHAGQWGTICDD 193

Qy        504 RWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKH 563
              :|   :| |:|:|||       :    :|. | |||| ||:| | ||| :|  |  | ||:|
Db        194 QWDNADADVICRQLGLSGIAKAWHQAHFGEGSGPILLDEVRCTGNELSIEQCPKSSWGEH 253

Qy        564 NCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQ 623
              || |:||  |:|       :|| || :   ||||||::|:|||||||||       | |
Db        254 NCGHKEDAGVSCVPLTDGVIRLAGGKSTHEGRLEVYYKGQWGTVCDDGWTEMNTYVAC-- 311

Qy        624 LDCPSSIIGMGLGNAST-----GYGK-IWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
                    ::|   |  |:     |  : |||||||| | |     |    || :|||| ||
Db        312 -----RLLGFKYGKQSSVNHFDGSNRPIWLDDVSCSGKEVSFIQCSRRQWGRHDCSHRED 366

Qy        678 VGVICSDASD------MELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQ 730
              ||: |   ||         :||| | ::   |:||| | |  | :| :||     | |:|||
Db        367 VGLTCYPDSDGHRLSPGFPIRLVDGENKKEGRVEVFVNGQWGTICDDGWTDKHAAVICRQ 426

Qy        731 LECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICS 790
              |      |     :| |    | | |  ||| | :| ||:: :  :  |   : :| :||
Db        427 LGYKGPARARTMAYFGEGKGPIHMDNVKCTGNEKALADCVKQDIGRHNCRHSEDAGVICD 486

Qy        791 -----------------------AHRQPRLVGADMPCSG--------RVEVKHADTWRS 818
                                        || |::| :    |        |:   | |  |
Db        487 YLEKKASSSGNKEMLSSGCGLRLLHRRQKRIIGGNNSLRGAWPWQASLRLRSAHGD-GRL 545

Qy        819 VC-----DSDFSLHAANVLCRELNCGDAISLSVGDH-------FGKGNG---LTWAEKFQ 863
              :|      | : | ||:   |  |   : :: |||:       | :  |   :      ::
Db        546 LCGATLLSSCWVLTAAHCFKRYGNNSRSYAVRVGDYHTLVPEEFEQEIGVQQIVIHRNYR 605

Qy        864 CEGSETHLALCPIVQHPEDTCIH-SREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWG 922
               : |:  :||  : | | : |   |  |   |      |     ||      :: |  ||
Db        606 PDRSDYDIALVRL-QGPGEQCARLSTHVLPACLPLWRERPQKTASNC------HITG-WG 657

Qy        923 SLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN---ESLLDNC 979
                       :  |   | |    |:      :: ||   ::  |  | ||    :: :|:|
Db        658 ---------DTGRAYSRTLQ-QAAVPLLPKRFCKERYKGLFTGRMLCAGNLQEDNRVDSC 707

Qy        980 Q 980
              |
Db        708 Q 708


RESULT 4
CD5L_HUMAN
ID   CD5L_HUMAN     STANDARD;      PRT;   347 AA.
AC   O43866;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   CD5 antigen-like precursor (SP-alpha) (CT-2) (IgM-associated peptide).
GN   CD5L.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF N-TERMINUS, AND TISSUE SPECIFICITY.
RC   TISSUE=Spleen;
RX   MEDLINE=97197777; PubMed=9045627;
RA   Gebe J.A., Kiener P.A., Ring H.Z., Li X., Francke U., Aruffo A.;
RT   "Molecular cloning, mapping to human chromosome 1 q21-q23, and cell
```

```
RT   binding characteristics of Spalpha, a new member of the scavenger
RT   receptor cysteine-rich (SRCR) family of proteins.";
RL   J. Biol. Chem. 272:6151-6158(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Miyazaki T., Yusa S.;
RT   "Human CT-2 cDNA.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Grafham D.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 20-31, AND BINDING TO IGM.
RX   MEDLINE=94308565; PubMed=8034987;
RA   Tissot J.D., Schifferli J.A., Hochstrasser D.F., Pasquali C.,
RA   Spertini F., Clement F., Frutiger S., Paquet N., Hughes G.J.,
RA   Schneider P.;
RT   "Two-dimensional polyacrylamide gel electrophoresis analysis of
RT   cryoglobulins and identification of an IgM-associated peptide.";
RL   J. Immunol. Methods 173:63-75(1994).
RN   [5]
RP   PARTIAL SEQUENCE, AND MASS SPECTROMETRY.
RX   PubMed=11981870;
RA   Tissot J.D., Sanchez J.C., Vuadens F., Scherl A., Schifferli J.A.,
RA   Hochstrasser D.F., Schneider P., Duchosal M.A.;
RT   "IgM are associated to Spalpha (CD5 antigen-like).";
RL   Electrophoresis 23:1203-1206(2002).
CC   -!- FUNCTION: May play a role in the regulation of the immune
CC       system. Seems to play a role as an inhibitor of apoptosis.
CC   -!- SUBUNIT: Associated with IgM.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed in spleen, lymph node, thymus, bone
CC       marrow, and fetal liver, but not in nonlymphoid tissues.
CC   -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U82812; AAB91989.1; -.
DR   EMBL; AF011429; AAD01446.1; -.
DR   EMBL; AL139409; CAC19458.1; -.
DR   Genew; HGNC:1690; CD5L.
DR   MIM; 602592; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 3.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 3.
DR   PROSITE; PS00420; SRCR_1; 2.
DR   PROSITE; PS50287; SRCR_2; 3.
KW   Signal; Repeat; Apoptosis.
FT   SIGNAL        1     19
FT   CHAIN        20    347       CD5 ANTIGEN-LIKE.
FT   DOMAIN       24    125       SRCR 1.
FT   DOMAIN      138    239       SRCR 2.
FT   DOMAIN      244    346       SRCR 3.
SQ   SEQUENCE   347 AA;  38088 MW;  40A8BE08F9495D83 CRC64;

  Query Match             8.5%;  Score 691.5;  DB 1;  Length 347;
  Best Local Similarity   41.3%;  Pred. No. 1.8e-41;
  Matches  148;  Conservative   40;  Mismatches  139;  Indels   31;  Gaps    9;

Qy         21 LFSAVVTCILLLNSC----FLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDG 76
              |||       |:|  |     || |           :|||  |   | | |||: :|||||||||
Db          4 LFS------LILAICTRPGFLASPSG-----VRLVGGLHRCEGRVEVEQKGQWGTVCDDG 52

Qy         77 WNTTASTVVCKQLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWG 133
              |:     |:|::|||   :        :     :  |:  :   ||| |  |   |  :|:  |
Db         53 WDIKDVAVLCRELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEE-- 110

Qy        134 SHNCYHGEDVGVNC------YGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLN 187
               ::| | || | :|          :         |:|| ||    | |||||| | :| |:|   ||:|
Db        111 VYDCSHDEDAGASCENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLR 170

Qy        188 TAAVVCRQLGCPSSFISSGVVNSPAVLR-PIWLDDILCQGNELALWNCRHRGWGNHDCSH 246
               | ||||||||   : ::       |    |   |  ||||    :  |  |  |   |  :|      || : |:|
Db        171 AAKVVCRQLGCGRAVLTQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNH 230

Qy        247 NEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCG 306
              :||    :  |  |   |  ||||||  |  |  ||:|:   :|  ||:||    |          ||||||||
Db        231 DEDTWVECEDPFD--LRLVGGDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCG 288

Qy        307 TAL--HFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS 362
               :|    |               |    :||| |  |||  |   |   |:|         || ||  ||:||||
Db        289 KSLSPSFRDRKCYGPGVGRIWLDNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICS 346


RESULT 5
CD5L_MOUSE
ID   CD5L_MOUSE     STANDARD;      PRT;   352 AA.
AC   Q9QWK4; O35300; O35301; Q91W05;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   CD5 antigen-like precursor (SP-alpha) (CT-2) (Apoptosis inhibitory 6)
DE   (Apoptosis inhibitor expressed by macrophages).
GN   CD5L OR API6 OR AIM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., VARIANTS MET-61; ASN-197 AND ARG-205, TISSUE
RP   SPECIFICITY, AND GLYCOSYLATION.
RC   STRAIN=BALB/c; TISSUE=Thymus;
RX   MEDLINE=20118142; PubMed=10651944;
RA   Gebe J.A., Llewellyn M.-B.C., Hoggatt H., Aruffo A.;
RT   "Molecular cloning, genomic organization and cell-binding
RT   characteristics of mouse Spalpha.";
RL   Immunology 99:78-86(2000).
RN   [2]
RP   SEQUENCE FROM N.A., VARIANT ASN-197, AND FUNCTION IN APOPTOSIS.
RX   MEDLINE=99110826; PubMed=9892623;
RA   Miyazaki T., Hirokami Y., Matsuhashi N., Takatsuka H., Naito M.;
RT   "Increased susceptibility of thymocytes to apoptosis in mice lacking
RT   AIM, a novel murine macrophage-derived soluble factor belonging to
RT   the scavenger receptor cysteine-rich domain superfamily.";
RL   J. Exp. Med. 189:413-422(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May play a role in the regulation of the immune
CC       system. Seems to play a role as an inhibitor of apoptosis.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed in thymus, liver, spleen and lymph
CC       nodes.
CC   -!- PTM: Glycosylated.
CC   -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF018268; AAB70571.1; -.
DR   EMBL; AF018269; AAB70572.1; -.
```

```
DR   EMBL; AF011428; AAD01445.1; -.
DR   EMBL; BC006799; AAH06799.1; -.
DR   MGD; MGI:1333862; Api6.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 3.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 3.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 3.
KW   Signal; Repeat; Glycoprotein; Apoptosis; Polymorphism.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    352       CD5 ANTIGEN-LIKE.
FT   DOMAIN       27    128       SRCR 1.
FT   DOMAIN      141    241       SRCR 2.
FT   DOMAIN      246    348       SRCR 3.
FT   CARBOHYD     99     99       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    229    229       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    316    316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      61     61       V -> M.
FT   VARIANT     197    197       S -> N.
FT   VARIANT     205    205       W -> R.
FT   CONFLICT     13     13       S -> N (IN REF. 1; AAB70571).
FT   CONFLICT    113    113       D -> Y (IN REF. 2).
SQ   SEQUENCE   352 AA;  38835 MW;  414984AACEF13B23 CRC64;

  Query Match             8.2%;  Score 664.5;  DB 1;  Length 352;
  Best Local Similarity   39.1%;  Pred. No. 1.5e-39;
  Matches  131;  Conservative   55;  Mismatches  134;  Indels   15;  Gaps    9;

Qy    253 TCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHF- 311
           :|:  |   :::||||  :|| ||||::  |:|||||    |:     |||::| ||   :
Db     18 SCFSESPTKVQLVGGAHRCEGRVEVEHNGQWGTVCDDGWDRRDVAVVCRELNCGAVIQTP 77

Qy    312 AGLPHLQSGSDV-VWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSD-GADL-- 367
           |  :     |:  |  | : || |.:| |   |   |   :  | ||| |: |    | :  :||
Db     78 RGASYQPPASEQRVLIQGVDCNGTEDTLAQCELNYDV-FDCSHEEDAGAQCENPDSDLLF 136

Qy    368 ---ELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKP 424
              ::|| ||   :| |||||     || |:|    |  : :  |||:||||  ::
Db    137 IPEDVRLVDGPGHCQGRVEVLHQSQWSTVCKAGWNLQVSKVVCRQLGCGRALLTYGSCNK 196

Qy    425 SNEAR-DIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHS 483
            | : :  ||:  :||:| |: | |      : :    |   |  : |  |    :|:||| :
Db    197 STQGKGPIWMGKMSCSGQEANLRSCLL-SRLENNCTHGEDTWMECED--PFELKLVGGDT 253

Qy    484 PCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMT--YFKEASGPIWLD 541
           || ||||| ::| ||:|| | |   :    |||||||||| :|      |    :    :| ||||
Db    254 PCSGRLEVLHKGSWGSVCDDNWGEKEDQVVCKQLGCGKSLHPSPKTRKIYGPGAGRIWLD 313

Qy    542 DVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCS 576
           ||:| | | ::  | |  || |:| |:||| | |:
Db    314 DVNCSGKEQSLEFCRHRLWGYHDCTHKEDVEVICT 348

RESULT 6
LOL4_MOUSE
ID   LOL4_MOUSE     STANDARD;      PRT;   757 AA.
AC   Q924C6;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lysyl oxidase homolog 4 precursor (EC 1.4.3.-) (Lysyl oxidase-like
DE   protein 4) (Lysyl oxidase related protein C).
GN   LOXL4 OR LOXC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21316447; PubMed=11292829;
RA   Ito H., Akiyama H., Iguchi H., Iyama Ki K., Miyamoto M., Ohsawa K.,
RA   Nakamura T.;
RT   "Molecular cloning and biological activity of a novel lysyl oxidase-
RT   related gene expressed in cartilage.";
RL   J. Biol. Chem. 276:24023-24029(2001).
CC   -!- FUNCTION: May modulate the formation of a collagenous
CC       extracellular matrix.
CC   -!- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (Potential).
CC   -!- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
CC       A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
CC       LYSINE (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF338440; AAK71933.1; -.
DR   MGD; MGI:1914823; Loxl4.
DR   InterPro; IPR001695; Lysyl_oxidase.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF01186; Lysyl_oxidase; 1.
DR   Pfam; PF00530; SRCR; 4.
DR   ProDom; PD013887; Lysyl_oxidase; 1.
DR   PROSITE; PS00926; LYSYL_OXIDASE; FALSE_NEG.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 4.
KW   Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    757       LYSYL OXIDASE HOMOLOG 4.
FT   DOMAIN       33    134       SRCR 1.
FT   DOMAIN      160    288       SRCR 2.
FT   DOMAIN      312    412       SRCR 3.
FT   DOMAIN      422    530       SRCR 4.
FT   DOMAIN      534    737       LYSYL-OXIDASE LIKE.
FT   METAL       612    612       COPPER (POTENTIAL).
FT   METAL       614    614       COPPER (POTENTIAL).
FT   METAL       616    616       COPPER (POTENTIAL).
FT   MOD_RES     639    639       CROSS-LINKED TO TYROSYLQUINONE (BY
FT                                SIMILARITY).
FT   MOD_RES     675    675       TYROSYLQUINONE (BY SIMILARITY).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    630    630       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   757 AA;  84705 MW;  D9861368F63B7B2A CRC64;

  Query Match             7.8%;  Score 633.5;  DB 1;  Length 757;
  Best Local Similarity   32.0%;  Pred. No. 5.2e-37;
  Matches  165;  Conservative   67;  Mismatches  190;  Indels   93;  Gaps   17;

Qy     42 NGTDLELRLVN-GDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM--- 97
          :||  :||||    |  |   |  :||  |||||||||  :      :|| |:|||   |   |:
Db     28 SGTK-KLRLVGPADRPEEGRLEVLHQGQWGTVCDDDFALQEATVACRQLG--FESALTWA 84

Qy     98 --FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGL 155
             ::||     | ||||:| |  |  |  |  | :|     ||   :|  |  |||||  |:     |
Db     85 HSAKYGQG---EGPIWLDNVRCLGTEKTLDQCGSNGWGISDCRHSEDVGVVCHPRRQHGY 141

Qy    156 RLVDGNNS------------------------CSGRVEVKFQERWGTICDDGWNLNTA 189
             :|:                             | |||::    |    :||  || :| :
Db    142 HSEKVSNALGPQGRRLEEVRLKPILASAKRHSPVTEGAVEVRYDGHWRQVCDQGWTMNNS 201

Qy    190 AVVCRQLGCPS-SFISS--------------GVVNSPAVLRPIWLDDILCQGNELALWN 233
          |||  || ||  : ::|                   :||          |:   :  |  |  |  |
Db    202 RVVCGMLGFPSQTSVNSHYYRKVWNLKMKDPKSRLNSLTKKNSFWIHRVDCFGTEPHLAK 261

Qy    234 CR-----HRGWGNHDCSHNEDVTLTCY-----------------DSSDLELRLVGGTNRC 271
          |:      ||       |      ::|                   : :|::|| |
```

```
Db     262 CQVQVAPGRGKLRPACPGGMHAVVSCVAGPHFRRQKPKPTRKESHAEELKVRLRSGAQVG 321

Qy     272 MGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSC 331
           ||||: :   :||||| |:||  :| |||:||| |:|           |  |    : |  | |
Db     322 EGRVEVLMNRQWGTVCDHRWNLISASVVCRQLGFGSAREALFGAQLGQGLGPIHLSEVRC 381

Qy     332 SGNESFLWDC-RHSGTVNFDCLHQNDVSVICS---DGADLELRLADGSNNCSGRVEVRIH 387
           | |  | ||     |: |  | | || :| |:     |    ::||| | |:  | |||::
Db     382 RGYERTLGDCLALEGSQN-GCQHANDAAVRCNIPDMGFQNKVRLAGGRNSEEGVVEVQVE 440

Qy     388 ----EQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA---KPSNEARDIWINSISCTG 440
               :| |:|  :|    :|:| |:|||  |: |   :        : : ||::: :: : |:|
Db     441 VNGGPRWGTVCSDHWGLTEAMVTCRQLGLGFANFALKDTWYWQGTPEAKEVVMSGVRCSG 500

Qy     441 NESALWDCTYDGKAKRTCFR---RSDAGVICSDKA 472
           | ||  |   |    |     |  ||| | : |
Db     501 TEMALQQCQRHGPVH--CSHGPGRFSAGVACMNSA 533
```

RESULT 7
LOL3_MOUSE

```
ID   LOL3_MOUSE     STANDARD;      PRT;   754 AA.
AC   Q9Z175; Q9JJ39;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lysyl oxidase homolog 3 precursor (EC 1.4.3.-) (Lysyl oxidase-like
DE   protein 3) (Lysyl oxidase related protein 2).
GN   LOXL3 OR LOR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J, and 129/SvJ; TISSUE=Muscle;
RX   MEDLINE=99126643; PubMed=9927484;
RA   Jang W., Hua A., Spilson S.V., Miller W., Roe B.A., Meisler M.H.;
RT   "Comparative sequence of human and mouse BAC clones from the mnd2
RT   region of chromosome 2p13.";
RL   Genome Res. 9:53-61(1999).
CC   -!- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (Potential).
CC   -!- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
CC       A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
CC       LYSINE (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF053368; AAC83205.1; -.
DR   EMBL; AF084363; AAC95338.1; -.
DR   MGD; MGI:1337004; Loxl3.
DR   InterPro; IPR001695; Lysyl_oxidase.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 4.
DR   Pfam; PF01186; Lysyl_oxidase; 1.
DR   PRINTS; PR00074; LYSYLOXIDASE.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD013887; Lysyl_oxidase; 1.
DR   SMART; SM00202; SR; 4.
DR   PROSITE; PS00926; LYSYL_OXIDASE; 1.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 4.
KW   Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    754       LYSYL OXIDASE HOMOLOG 3.
FT   DOMAIN       45    146       SRCR 1.
FT   DOMAIN      170    283       SRCR 2.
FT   DOMAIN      308    408       SRCR 3.
FT   DOMAIN      418    526       SRCR 4.
FT   DOMAIN      530    733       LYSYL-OXIDASE LIKE.
FT   METAL       608    608       COPPER (POTENTIAL).
FT   METAL       610    610       COPPER (POTENTIAL).
FT   METAL       612    612       COPPER (POTENTIAL).
FT   MOD_RES     635    635       CROSS-LINKED TO TYROSYLQUINONE (BY
FT                                SIMILARITY).
FT   MOD_RES     671    671       TYROSYLQUINONE (BY SIMILARITY).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    391    391       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    482    482       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    626    626       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   754 AA;  83681 MW;  F8758400943F4FF2 CRC64;

  Query Match             7.7%;  Score 629;  DB 1;  Length 754;
  Best Local Similarity   31.0%;  Pred. No. 1.1e-36;
  Matches  155;  Conservative   72;  Mismatches  183;  Indels   90;  Gaps   16;

Qy      56 FCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKI 110
           | | ||::  |:|||:|||  :    |: |:|::||  |: |         ::|       |:|
Db      54 FYEGRVEIQRAGEWGTICDDDFTLQAAHVLCRELG--FTEATGWTHSAKYGPGT---GRI 108

Qy     111 WLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNC-------YGEANL---------- 153
           |||::|| | | :: ||  | ||: :| | || || |        : ::|:       .
Db     109 WLDNLSCRGTEGSVTECASRGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHQLQVE 168

Qy     154 ----------GLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS-- 201
                     | | :        | |||:  | |  :|| ||: : : |||  || |
Db     169 EVRLRPAVEWGRRPLP---VTEGLVEVRLPEGWSQVCDKGWSAHNSHVVCGMLGFPGEKR 225

Qy     202 -------FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHD---CSHNEDVT 251
                  ::    :|         | : | |  |   |   |     :  :|    ||
Db     226 VNMAFYRMLAQKKQHS------FGLHSVACVGTEAHLSLCSLEFYRANDTTRCSGGNPAV 279

Qy     252 LTCY-----------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNN 293
           ::|                  : :|| || ::  ||||:   | |||||  ||:
Db     280 VSCVLGPLYATFTGQKKQQHSKPQGEARVRLKGGAHQGEGRVEVLKAGTWGTVCDRKWDL 339

Qy     294 AAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLH 353
            || ||| :|| |||          :  | .  : |  | ||| |  ||  |            || |
Db     340 QAASVVCPELGFGTAREALSGARMGQGMGAIHLSEVRCSGQEPSLWRCPSKNITAEDCSH 399

Qy     354 QNDVSVICS---DGADLELRLADGSNNCSGRVEVRI----HEQWWTICDQNWKNEQALVV 406
             |  | |:   | : ::||: | :   |||||:|     | :|  ||  :|    :|:|
Db     400 SQDAGVRCNLPYTGVETKIRLSGGRSRYEGRVEVQIGIPGHLRWGLICGDDWGTLEAMVA 459

Qy     407 CKQLGCPFSVFGSRRA--KPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR---R 461
           |:|||  ::  | :         |     :: :: : |||:| :|  | :    :  || :   |
Db     460 CRQLGLGYANHGLQETWYWDSGNVTEVVMSGVRCTGSELSLNQCAHH-SSHITCKKTGTR 518

Qy     462 SDAGVICSDKADLDLRLVGA 481
             ||||||: |  || |  |
Db     519 FTAGVICSETAS-DLLLHSA 537
```

RESULT 8
LOL3_HUMAN

```
ID   LOL3_HUMAN     STANDARD;      PRT;   753 AA.
AC   P58215; Q96RS1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lysyl oxidase homolog 3 precursor (EC 1.4.3.-) (Lysyl oxidase-like
DE   protein 3).
GN   LOXL3 OR LOXL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21181571; PubMed=11284725;
RA   Maki J.M., Kivirikko K.I.;
RT   "Cloning and characterization of a fourth human lysyl oxidase
RT   isoenzyme.";
RL   Biochem. J. 355:381-387(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21280915; PubMed=11386757;
RA   Jourdan-Le Saux C., Tomshe A., Ujfalusi A., Jia L., Csiszar K.;
RT   "Central nervous system, uterus, heart, and leukocyte expression of
RT   the LOXL3 gene, encoding a novel lysyl oxidase-like protein.";
RL   Genomics 74:211-218(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21233589; PubMed=11334717;
RA   Huang Y., Dai J., Tang R., Zhao W., Zhou Z., Wang W., Ying K., Xie Y.,
RA   Mao Y.;
RT   "Cloning and characterization of a human lysyl oxidase-like 3 gene
RT   (hLOXL3).";
RL   Matrix Biol. 20:153-157(2001).
CC   -!- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (Potential).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES, THE HIGHEST LEVELS
CC       AMONG THE TISSUES STUDIED BEING SEEN IN THE PLACENTA, HEART,
CC       OVARY, TESTIS, SMALL INTESTINE AND SPLEEN.
CC   -!- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
CC       A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
CC       LYSINE (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF282619; AAK51671.1; -.
DR   EMBL; AF311313; AAK63205.1; -.
DR   EMBL; AF284815; AAK91134.1; -.
DR   Genew; HGNC:13869; LOXL3.
DR   InterPro; IPR001695; Lysyl_oxidase.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 4.
DR   Pfam; PF01186; Lysyl_oxidase; 1.
DR   PRINTS; PR00074; LYSYLOXIDASE.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD013887; Lysyl_oxidase; 1.
DR   SMART; SM00202; SR; 4.
DR   PROSITE; PS00926; LYSYL_OXIDASE; 1.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 4.
KW   Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    753       LYSYL OXIDASE HOMOLOG 3.
FT   DOMAIN       44    145       SRCR 1.
FT   DOMAIN      169    282       SRCR 2.
FT   DOMAIN      307    407       SRCR 3.
FT   DOMAIN      417    525       SRCR 4.
FT   DOMAIN      529    732       LYSYL-OXIDASE LIKE.
FT   METAL       607    607       COPPER (POTENTIAL).
FT   METAL       609    609       COPPER (POTENTIAL).
FT   METAL       611    611       COPPER (POTENTIAL).
FT   MOD_RES     634    634       CROSS-LINKED TO TYROSYLQUINONE (BY
FT                                SIMILARITY).
FT   MOD_RES     670    670       TYROSYLQUINONE (BY SIMILARITY).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    266    266       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    390    390       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    481    481       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    625    625       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    159    159       E -> K (IN REF. 3).
SQ   SEQUENCE   753 AA;  83166 MW;  582C46DA25E05A69 CRC64;

  Query Match             7.6%;  Score 622.5;  DB 1;  Length 753;
  Best Local Similarity   31.2%;  Pred. No. 3.1e-36;
  Matches  167;  Conservative   71;  Mismatches  202;  Indels   95;  Gaps   19;

Qy         30 LLLNSCFLISSFNGT-------------DLELRLVN-GDGPCSGTVEVKFQGQWGTVCD 74
              |||  | |  ||  |:             |   ||        |  | ||::  |:|||:||
Db         14 LLL--CLLCSSCLGSPSPSTGPEKKAGSQGLRFRLAGFPRKPYEGRVEIQRAGEWGTICD 71

Qy         75 DGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECQH 129
              | :   |: ::|::||  |: |         ::|       |:||||::||  | |  :: ||
Db         72 DDFTLQAAHILCRELG--FTEATGWTHSAKYGPGT---GRIWLDNLSCSGTEQSVTECAS 126

Qy        130 REWGSHNCYHGEDVGVNC-------YGEANL-------------------GLRLVDGNN 162
              | ||: :| | || || |         : ::|:                       | | :
Db        127 RGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHHLQVEEVRIRPAVGWGRRPLP--- 183

Qy        163 SCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLR---PIWL 219
                  | |||:   : |    :|| ||:  :  : |||   || ||                 |      |
Db        184 VTEGLVEVRLPDGWSQVCDKGWSAHNSHVVCGMLGFPSEKRVNAAFYRLLAQRQQHSFGL 243

Qy        220 DDILCQGNELALWNCRHRGWGNHD---CSHNEDVTLTC-----YDSS------------ 258
                 : | | |   |   |    :  :|   |          ::|     | :|
Db        244 HGVACVGTEAHLSLCSLEFYRANDTARCPGGGPAVVSCVPGPVYAASSGQKKQQQSKPQG 303

Qy        259 DLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
               :  :|| || :     ||||:     |||||  ||:  || |||::|| |:|          :
Db        304 EARVRLKGGAHPGEGRVEVLKASTWGTVCDRKWDLHAASVVCRELGFGSAREALSGARMG 363

Qy        319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS---DGADLELRLADGS 375
                |    : |   | |||  |   || |  |            || |   |  | |:      ||:  :||: |
Db        364 QGMGAIHLSEVRCSGQELSLWKCPHKNITAEDCSHSQDAGVRCNLPYTGAETRIRLSGGR 423

Qy        376 NNCSGRVEVRIHE----QWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA--KPSNEAR 429
                :    |||||:|         :|  ||  :|     :|:| |:|||   ::   | :             |
Db        424 SQHEGRVEVQIGGPGPLRWGLICGDDWGTLEAMVACRQLGLGYANHGLQETWYWDSGNIT 483

Qy        430 DIWINSISCTGNESALWDCTYDGKAKRTCFR---RSDAGVICSDKADLDLRLVGA 481
              :: ::  : |||.| :|   | : |        ||  |     ||||||: |  ||  |   |
Db        484 EVVMSGVRCTGTELSLDQCAHHG-THITCKRTGTRFTAGVICSETAS-DLLLHSA 536


RESULT 9
LOL4_HUMAN
ID   LOL4_HUMAN     STANDARD;      PRT;   756 AA.
AC   Q96JB6; Q96PC0; Q96DY1;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lysyl oxidase homolog 4 precursor (EC 1.4.3.-) (Lysyl oxidase-like
DE   protein 4) (Lysyl oxidase related protein C).
GN   LOXL4 OR LOXC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=21316447; PubMed=11292829;
RA   Ito H., Akiyama H., Iguchi H., Iyama K., Miyamoto M., Ohsawa K.,
RA   Nakamura T.;
RT   "Molecular cloning and biological activity of a novel lysyl oxidase-
RT   related gene expressed in cartilage.";
RL   J. Biol. Chem. 276:24023-24029(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21550107; PubMed=11691589;
```

```
RA   Maki J.M., Tikkanen H., Kivirikko K.I.;
RT   "Cloning and characterization of a fifth human lysyl oxidase
RT   isoenzyme: the third member of the lysyl oxidase-related subfamily
RT   with four scavenger receptor cysteine-rich domains.";
RL   Matrix Biol. 20:493-496(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Asuncion L.P., Fogelgren B., Fong K.S.K., Fong S.F.T., Kim Y.,
RA   Csiszar K.;
RT   "A novel human lysyl oxidase-like gene (LOXL4) on chromosome 10q24 has
RT   an altered SRCR domain and is down-regulated by H-ras.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May modulate the formation of a collagenous
CC       extracellular matrix.
CC   -!- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (Potential).
CC   -!- TISSUE SPECIFICITY: Expressed in many tissues, the highest levels
CC       among the tissues studied being in the skeletal muscle, testis and
CC       pancreas. Expressed in cartilage.
CC   -!- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
CC       A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
CC       LYSINE (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF338441; AAK71934.1; -.
DR   EMBL; AY036093; AAK64186.1; -.
DR   EMBL; AF395336; AAL27543.1; -.
DR   EMBL; BC013153; AAH13153.1; -.
DR   Genew; HGNC:17171; LOXL4.
DR   InterPro; IPR001695; Lysyl_oxidase.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF01186; Lysyl_oxidase; 1.
DR   Pfam; PF00530; SRCR; 4.
DR   ProDom; PD013887; Lysyl_oxidase; 1.
DR   PROSITE; PS00926; LYSYL_OXIDASE; FALSE_NEG.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 4.
KW   Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT   SIGNAL         1     24       POTENTIAL.
FT   CHAIN         25    756       LYSYL OXIDASE HOMOLOG 4.
FT   DOMAIN        32    133       SRCR 1.
FT   DOMAIN       159    287       SRCR 2.
FT   DOMAIN       311    411       SRCR 3.
FT   DOMAIN       421    529       SRCR 4.
FT   DOMAIN       533    736       LYSYL-OXIDASE LIKE.
FT   METAL        611    611       COPPER (POTENTIAL).
FT   METAL        613    613       COPPER (POTENTIAL).
FT   METAL        615    615       COPPER (POTENTIAL).
FT   MOD_RES      638    638       CROSS-LINKED TO TYROSYLQUINONE (BY
FT                                 SIMILARITY).
FT   MOD_RES      674    674       TYROSYLQUINONE (BY SIMILARITY).
FT   CARBOHYD     198    198       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     629    629       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT       3      3       W -> R (IN REF. 4).
FT   CONFLICT     101    101       R -> Q (IN REF. 4).
FT   CONFLICT     405    405       D -> A (IN REF. 4).
FT   CONFLICT     493    493       S -> G (IN REF. 3).
FT   CONFLICT     539    539       A -> T (IN REF. 3).
FT   CONFLICT     542    542       V -> A (IN REF. 3).
FT   CONFLICT     703    703       Y -> H (IN REF. 3).
SQ   SEQUENCE   756 AA;  84483 MW;  13051ACADB922BBC CRC64;

  Query Match             7.6%;  Score 618;  DB 1;  Length 756;
  Best Local Similarity   31.2%;  Pred. No. 6.4e-36;
  Matches  159;  Conservative   68;  Mismatches  191;  Indels   92;  Gaps   16;

Qy         47 ELRLVNGDG-PCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM-----FRF 100
              :||||  :  |  | :||  |||||||||| :     :|| |:|||  |  |:      ::
Db         31 KLRLVGPESKPEEGRLEVLHQGQWGTVCDDNFAIQEATVACRQLG--FEAALTWAHSAKY 88

Qy        101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
              ||     | ||||:| | | ||:| :|     ||  :| | ||||| |:    :  |
Db         89 GQG---EGPIWLDNVRCVGTESSLDQCGSNGWGVSDCSHSEDVGVICHPRRHRGYLSETV 145

Qy        161 NNS-------------------------CSGRVEVKFQERWGTICDDGWNLNTAAVVCR 194
              :|:                         | ||||::  |   :|| || :| : |||
Db        146 SNALGPQGRRLEEVRLKPILASAKQHSPVTEGAVEVKYEGHWRQVCDQGWTMNNSRVVCG 205

Qy        195 QLGCPSSF-ISS---------------GVVNSPAVLRPIWLDDILCQGNELALWNCR--- 235
              || ||   :  |                     :  |       |:  :  | | |  : ||:
Db        206 NLGFPSEVPVDSHYYRKVWDLKMRDPKSRLKSLTNKNSFWIHQVTCLGTEPHMANCQVQV 265

Qy        236 --HRGWGNHDCSHNEDVTLTCYD-----------------SSDLELRLVGGTNRCMGRVE 276
                 ||      |         ::|                 : :   :||   |      ||||
Db        266 APARGKLRPACPGGMHAVVSCVAGPHFRPPKTKPQRKGSWAEEPRVRLRSGAQVGEGRVE 325

Qy        277 LKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNES 336
              : :   :||||| |:||   :| |||:||| |:|          |  |    :  |  | | | |
Db        326 VLMNRQWGTVCDHRWNLISASVVCRQLGFGSAREALFGARLGQGLGPIHLSEVRCRGYER 385

Qy        337 FLWDC-RHSGTVNFDCLHQNDVSVIC---SDGADLELRLADGSNNCSGRVEVRIH----E 388
               | ||      |:  |   | |:|| :| |      :  |      ::|||  |        | :||::
Db        386 TLSDCPALEGSQN-GCQHENDAAVRCNVPNMGFQNQVRLAGGRIPEEGLLEVQVEVNGVP 444

Qy        389 QWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA---KPSNEARDIWINSISCTGNESAL 445
              :| ::| :||   :|:| |:|||  |::          :     |:::  ::  : |:| | ||
Db        445 RWGSVCSENWGLTEAMVACRQLGLGFAIHAYKETWFWSGTPRAQEVVMSGVRCSGTELAL 504

Qy        446 WDCTYDGKAKRTCFR---RSDAGVICSDKA 472
                  |    |      |      |   ||| | | |
Db        505 QQCQRHGPVH--CSHGGGRFLAGVSCMDSA 532

RESULT 10
CD6_MOUSE
ID   CD6_MOUSE      STANDARD;      PRT;   665 AA.
AC   Q61003; Q61004; Q60679;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   T-cell differentiation antigen CD6 precursor.
GN   CD6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Thymocytes;
RX   MEDLINE=96062291; PubMed=7594475;
RA   Robinson W.H., Prohaska S.S., Santoro J.C., Robinson H.L.,
RA   Parnes J.R.;
RT   "Identification of a mouse protein homologous to the human CD6 T cell
RT   surface protein and sequence of the corresponding cDNA.";
RL   J. Immunol. 155:4739-4748(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RX   MEDLINE=95174761; PubMed=7870060;
RA   Whitney G., Bowen M., Neubauer M., Aruffo A.;
RT   "Cloning and characterization of murine CD6.";
RL   Mol. Immunol. 32:89-92(1995).
```

```
CC   -!- FUNCTION: INVOLVED IN CELL ADHESION. BINDS TO CD166.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2 AND
CC       3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THYMUS, LYMPH NODE,
CC       AND SPLEEN.
CC   -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U37543; AAA81383.1; -.
DR   EMBL; U37544; AAA81384.1; -.
DR   EMBL; U12434; AAA64867.1; -.
DR   MGD; MGI:103566; Cd6.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 3.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 3.
DR   PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR   PROSITE; PS50287; SRCR_2; 3.
KW   Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    665       T-CELL DIFFERENTIATION ANTIGEN CD6.
FT   DOMAIN       17    398       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    399    419       POTENTIAL.
FT   DOMAIN      420    665       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       45    155       SRCR 1.
FT   DOMAIN      160    259       SRCR 2.
FT   DOMAIN      264    360       SRCR 3.
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     34     34       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    117    117       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    228    228       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    344    344       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    367    367       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    461    500       APMLFIQPRVPADSDSSSDSDYEHYDFSSQPPVALTTFYN
FT                                -> D (IN ISOFORM 3).
FT   VARSPLIC    557    586       NSDSSTSSEEGYCNDPSSKPPPWNSQAFYS -> KDKASGV
FT                                RAESWVEQTGSGHFLGVVKGHAG (IN ISOFORM 2).
FT   VARSPLIC    587    665       MISSING (IN ISOFORM 2).
FT   CONFLICT      5      5       L -> F (IN REF. 2).
FT   CONFLICT      8      8       A -> T (IN REF. 2).
FT   CONFLICT    194    194       D -> A (IN REF. 2).
FT   CONFLICT    316    317       AV -> GL (IN REF. 2).
SQ   SEQUENCE   665 AA;  72299 MW;  379675278C2CAB53 CRC64;

  Query Match             7.6%;  Score 614.5;  DB 1;  Length 665;
  Best Local Similarity   28.6%;  Pred. No. 9.8e-36;
  Matches  162;  Conservative   57;  Mismatches  154;  Indels  193;  Gaps   14;

Qy    900 VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERS 959
          :||||| |:| | |:: :|   |  :|  ||:        :|: |:||
Db     45 IRLVNGSSRCSGSVKV-LLESWEPVCAAHWNRAATEAVCKALNCG--------------- 88

Qy    960 VRVWGHRFHCLGNESLLDNCQMTVLGAPP------CIHGNTVSVICTGSLTQPLFPC--- 1010
                         |: ::| |   |           ||| |    |      |   |
Db     89 -----------------DSGKVTYLMPPTSELPPGATSGNTSSAGNTTWARAPTERCRGA 131

Qy   1011 ---LANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDL 1067
              | |   |:          : | |:: :||||| |||||||:   | |||:||| |||
Db    132 NWQFCKVQDQECSSDRRLVWVTCAENQAVRLVDGSSRCAGRVEMLEHGEWGTVCDDTWDL 191

Qy   1068 SDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRH 1127
            | ||||::| || |  |      ||  | |||   | :||:|||::|| || |     |: | |
Db    192 QDDHVVCKQLKCGWAVKALAGLHFTPGQGPIHRDQVNCSGTEAYLWDCPGRPGDQY-CGH 250

Qy   1128 KEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCG 1187
          ||||||:|||
Db    251 KEDAGVVCSE-------------------------------------------------- 260

Qy   1188 ENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIR 1247
                                                             ::|       |
Db    261 --------------------------------------------------HQSW-------R 265

Qy   1248 VRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGT 1307
          : ||    | |:||::  | | ||||  |   :||:|:|: ||||||:|   |
Db    266 LTGGIDSCEGQVEVYFRGVWSTVCDSEWYPSEAKVLCRSLGCGSAVARPRAVPHS----- 320

Qy   1308 IWLDD---MRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASS------- 1357
            ||        ||| |   |   |   :       : |         | | ||| | : || |:
Db    321 --LDGRMYYSCKGQEPALSTCSWRFNNSNLCSQSRAARVVCSG-SQRHLNLSTSEVPSRV 377

Qy   1358 -----------------GHLALILSSIFGLLLLV--LFILFLTWCRVQKQKHLPLRVST 1397
                           |    |||   : |:||||   :||: |     | :  |   ||   |:
Db    378 PVTIESSVPVSVKDKDSQGLTLLILCIVLGILLLVSTIFIVILL-LRAKGQYALPASVNH 436

Qy   1398 RRRGSLEE---NLFHEMETCLKREDP 1420
           :: :  :     | :| :    : :| |
Db    437 QQLSTANQAGINNYHPVPITIAKEAP 462


RESULT 11
LOL2_HUMAN
ID   LOL2_HUMAN     STANDARD;      PRT;   774 AA.
AC   Q9Y4K0; Q9Y5Y8; Q9BW70;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lysyl oxidase homolog 2 precursor (EC 1.4.3.-) (Lysyl oxidase-like
DE   protein 2) (Lysyl oxidase related protein 2) (Lysyl oxidase-related
DE   protein WS9-14).
GN   LOXL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97236759; PubMed=9079631;
RA   Saito H., Papaconstantinou J., Sato H., Goldstein S.;
RT   "Regulation of a novel gene encoding a lysyl oxidase-related protein
RT   in cellular adhesion and senescence.";
RL   J. Biol. Chem. 272:8157-8160(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 202-774 FROM N.A.
RC   TISSUE=Spleen, and Placenta;
RX   MEDLINE=99230328; PubMed=10212285;
RA   Jourdan-Le Saux C., Tronecker H., Bogic L., Bryant-Greenwood G.D.,
RA   Boyd C.D., Csiszar K.;
RT   "The LOXL2 gene encodes a new lysyl oxidase-like protein and is
RT   expressed at high levels in reproductive tissues.";
RL   J. Biol. Chem. 274:12939-12944(1999).
CC   -!- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (Potential).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES. HIGHEST EXPRESSION
CC       IN REPRODUCTIVE TISSUES, PLACENTA, UTERUS AND PROSTATE.
CC   -!- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
CC       A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
CC       LYSINE (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U89942; AAB49697.1; -.
DR   EMBL; BC000594; AAH00594.1; -.
DR   EMBL; AF117949; AAD34343.1; -.
DR   Genew; HGNC:6666; LOXL2.
DR   MIM; 606663; -.
DR   InterPro; IPR001695; Lysyl_oxidase.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 4.
DR   Pfam; PF01186; Lysyl_oxidase; 1.
DR   PRINTS; PR00074; LYSYLOXIDASE.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD013887; Lysyl_oxidase; 1.
DR   SMART; SM00202; SR; 4.
DR   PROSITE; PS00926; LYSYL_OXIDASE; 1.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 4.
KW   Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    774       LYSYL OXIDASE HOMOLOG 2.
FT   DOMAIN       58    159       SRCR 1.
FT   DOMAIN      188    302       SRCR 2.
FT   DOMAIN      326    425       SRCR 3.
FT   DOMAIN      435    544       SRCR 4.
FT   DOMAIN      548    751       LYSYL-OXIDASE LIKE.
FT   METAL       626    626       COPPER (POTENTIAL).
FT   METAL       628    628       COPPER (POTENTIAL).
FT   METAL       630    630       COPPER (POTENTIAL).
FT   MOD_RES     653    653       CROSS-LINKED TO TYROSYLQUINONE (BY
FT                                SIMILARITY).
FT   MOD_RES     689    689       TYROSYLQUINONE (BY SIMILARITY).
FT   CARBOHYD    288    288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    455    455       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    644    644       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    295    295       L -> Q (IN REF. 3).
FT   CONFLICT    570    570       M -> L (IN REF. 2).
FT   CONFLICT    652    652       H -> Q (IN REF. 3).
FT   CONFLICT    746    746       C -> S (IN REF. 3).
FT   CONFLICT    774    774       Q -> PVKKPAWSTPVFRPHHIFHGTSPQQLSLNECHVPSP
FT                                SPAPTLSRPLQLCLSSGGKGPSHHSWGAAT (IN REF.
FT                                3).
SQ   SEQUENCE   774 AA;  86724 MW;  9DF5D25D4824BCCD CRC64;

  Query Match             7.5%;  Score 607.5;  DB 1;  Length 774;
  Best Local Similarity   27.9%;  Pred. No. 3.6e-35;
  Matches  196;  Conservative   86;  Mismatches  258;  Indels  163;  Gaps   28;

Qy    260 LELRLVGGTNR-CMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
Db     56 IQLRLAGQKRKHSEGRVEVYYDGQWGTVCDDDFSIHAAHVVCRELGYVEAKSWTASSSYG 115

Qy    319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSD----GADLELRLADG 374
Db    116 KGEGPIWLDNLHCTGNEATLAACTSNGWGVTDCKHTEDVGVVCSDKRIPGFKFDNSLINQ 175

Qy    375 SNNCS-----------------------GRVEVRIHEQWWTICDQNWKNEQALVVCKQLG 411
Db    176 IENLNIQVEDIRIRAILSTYRKRTPVMEGYVEVKEGKTWKQICDKHWTAKNSRVVCGMFG 235

Qy    412 CP---------FSVFGSRRAKPSNEARDIWINSISCTGNESALWDC------TYDGKAKR 456
Db    236 FPGERTYNTKVYKMFASRR------KQRYWPFSMDCTGTEAHISSCKLGPQVSLDPMKNV 289

Qy    457 TCFRRSDAGVIC----------------SDKADLDL-RLVGAHSPCYGRLEVKYQGEWGT 499
Db    290 TCENGLPAVVSCVPGQVFSPDGPSRFRKAYKPEQPLVRLRGGAYIGEGRVEVLKNGEWGT 349

Qy    500 VCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSG 559
Db    350 VCDDKWDLVSASVVCRELGFGSAKEAVTGSRLGQGIGPIHLNEIQCTGNEKSIIDCKFNA 409

Qy    560 WGKHNCVHREDVIVTCSGDATW---GLRLVGGSNRCSGRLEVYFQGR----WGTVCDDGW 612
Db    410 -ESQGCNHEEDAGVRCNTPAMGLQKKLRLNGGRNPYEGRVEVLVERNGSLVWGMVCGQNW 468

Qy    613 NSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIW-----------LDDVSCDGDESDLW 660
Db    469 GIVEAMVVCRQL---------GLGFASNAFQETWYWHGDVNSNKVVMSGVKCSGTELSLA 519

Qy    661 SCRNSGWGNNDCSHSEDV----------GVICSD-ASDMEL--RLVGGSSRCAGKVEVN 706
Db    520 HCRHDG---------EDVACPQGGVQYGAGVACSETAPDLVLNAEMVQQTTYLEDRPMFM 570

Qy    707 VQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMS---NSGCT--- 760
Db    571 LQ------CA-----------MEENCLSASAAQTDPTTGYRRLLRFSSQIHNNGQSDFR 612

Qy    761 ---GGEASLW-DCIR-WEWKQTACH---LNMEASLICSAHRQPRLVGADMPCSGRVEVKH 812
Db    613 PKNGRHAWIWHDCHRHYHSMEVFTHYDLLNLNGTKVAEGHKASFCL-EDTECEGDIQKNY 671

Qy    813 ADTWRSVC----DSDFSLHAANVLCRELNCG--DAISLSVGDH 849
Db    672 E------CANFGDQGITMGCWDMYRHDIDCQWVDITDVPPGDY 708


RESULT 12
CD6_HUMAN
ID   CD6_HUMAN      STANDARD;      PRT;   668 AA.
AC   P30203; Q9UMF2; Q9Y4L0; Q9Y4K9; Q9Y4K8; Q9Y4K7;
DT   01-APR-1993 (Rel. 25, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   T-cell differentiation antigen CD6 precursor (T12) (TP120).
GN   CD6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 1-468 FROM N.A.
RX   MEDLINE=92013796; PubMed=1919444;
RA   Aruffo A., Melnick M.B., Linsley P.S., Seed B.;
RT   "The lymphocyte glycoprotein CD6 contains a repeated domain structure
RT   characteristic of a new family of cell surface and secreted
RT   proteins.";
RL   J. Exp. Med. 174:949-952(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=97166061; PubMed=9013954;
RA   Bowen M.A., Whitney G.S., Neubauer M., Starling G.C., Palmer D.,
RA   Zhang J., Nowak N.J., Shows T.B., Aruffo A.;
RT   "Structure and chromosomal location of the human CD6 gene: detection
RT   of five human CD6 isoforms.";
RL   J. Immunol. 158:1149-1156(1997).
CC   -!- FUNCTION: INVOLVED IN CELL ADHESION. BINDS TO CD166.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; CD6A (SHOWN HERE), CD6B, CD6C,
CC       CD6D AND CD6E; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD6 entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd6.htm".
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X60992; CAA43306.1; -.
DR   EMBL; U66142; AAC51161.1; -.
DR   EMBL; U66143; AAC51162.1; -.
DR   EMBL; U66144; AAC51163.1; -.
DR   EMBL; U66145; AAC51164.1; -.
DR   EMBL; U66146; AAC51165.1; -.
DR   PIR; S26741; S26741.
DR   Genew; HGNC:1691; CD6.
DR   MIM; 186720; -.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 3.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 3.
DR   PROSITE; PS00420; SRCR_1; 1.
DR   PROSITE; PS50287; SRCR_2; 3.
KW   Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    668       T-CELL DIFFERENTIATION ANTIGEN CD6.
FT   DOMAIN       18    402       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    403    423       POTENTIAL.
FT   DOMAIN      424    668       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       45    156       SRCR 1.
FT   DOMAIN      161    260       SRCR 2.
FT   DOMAIN      265    361       SRCR 3.
FT   CARBOHYD     28     28       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    229    229       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    345    345       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    431    462       MISSING (IN ISOFORM CD6B, ISOFORM CD6C
FT                                AND ISOFORM CD6D).
FT   VARSPLIC    463    504       VFMLPIQVQAPPPEDSDSGSDSDYEHYDFSAQPPVALTTFY
FT                                N -> D (IN ISOFORM CD6C AND ISOFORM
FT                                CD6E).
FT   VARSPLIC    613    647       MISSING (IN ISOFORM CD6D AND ISOFORM
FT                                CD6E).
FT   CONFLICT    463    468       VFMLPI -> GPGPAP (IN REF. 1).
FT   CONFLICT    613    613       MISSING (IN REF. 2; AAC51162).
SQ   SEQUENCE   668 AA;  71828 MW;  EFB434608012C49E CRC64;

  Query Match             7.3%;  Score 592;  DB 1;  Length 668;
  Best Local Similarity   28.7%;  Pred. No. 3.7e-34;
  Matches  154;  Conservative   59;  Mismatches  140;  Indels  184;  Gaps   15;

Qy     900 VRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERS 959
           ||| || | | | ||: :  |   |   ||   |  :|| | || | : :
Db      45 VRLTNGSSSCSGTVEVRLEASWEPACGALWDSRAAEAVCRALGCGGAEAASQ-------- 96

Qy     960 VRVWGHRFHCL-------------GNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQP 1006
                     |             || |:  |   |: |||         :::|:|:  :
Db      97 ----------LAPPTPELPPPPAAGNTSVAANA--TLAGAP--------ALLCSGAEWR- 135

Qy    1007 LFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWD 1066
               | |  :     : : | |:: ||||||   ||||||:  | ||::||| ||
Db     136 ----LCEVVEHACRSDGRRARVTCAENRALRLVDGGGACAGRVEMLEHGEWGSVCDDTWD 191

Qy    1067 LSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCR 1126
           | ||||||::|||| |  |     ||  | |||  | :||:| |::|| ||    ||| |
Db     192 LEDAHVVCRQLGCGWAVQALPGLHFTPGRGPIHRDQVNCSGAEAYLWDCPGLP-GQHYCG 250

Qy    1127 HKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGC 1186
           |||||||:|||
Db     251 HKEDAGVVCSE------------------------------------------------- 261

Qy    1187 GENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRI 1246
                                                        ::|
Db     262 ---------------------------------------------HQSW------- 265

Qy    1247 RVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTG 1306
           |: ||   | |:||:   | | ||||  |  :||:|:|| ||||:|:   :          :|
Db     266 RLTGGADRCEGQVEVHFRGVWNTVCDSEWYPSEAKVLCQSLGCGTAVERPKGLPHSL-SG 324

Qy    1307 TIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCS-GQSLKSL------------ 1353
            ::     | | |  | :|  :    : |        | | ||  :|| :|
Db     325 RMYYS---CNGEELTLSNCSWRFNNSNLCSQSLAARVLCSASRSLHNLSTPEVPASVQTV 381

Qy    1354 -----------NASSGHLALILSSI-FGLLLL--VLFILFLTWCRVQKQKHLPLRVS 1396
                      |  |  | |:: ||  |:|||  ::|| |:   |:: :  ||: |:
Db     382 TIESSVTVKIENKESRELMLLIPSIVLGILLLGSLIFIAFIL-LRIKGKYALPVMVN 437

RESULT 13
SPER_STRPU
ID   SPER_STRPU     STANDARD;      PRT;   532 AA.
AC   P16264;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Egg peptide speract receptor precursor.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 477-489.
RX   MEDLINE=89184581; PubMed=2538832;
RA   Dangott L.J., Jordan J.E., Bellet R.A., Garbers D.L.;
RT   "Cloning of the mRNA for the protein that crosslinks to the egg
RT   peptide speract.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:2128-2132(1989).
CC   -!- FUNCTION: RECEPTOR FOR THE EGG PEPTIDE SPERACT.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; J04518; AAA30078.1; -.
DR   PIR; A32751; A32751.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 4.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00202; SR; 4.
DR   PROSITE; PS00420; SRCR_1; 4.
DR   PROSITE; PS50287; SRCR_2; 4.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT   SIGNAL        1     30
FT   CHAIN        31    532       EGG PEPTIDE SPERACT RECEPTOR.
FT   DOMAIN       31    491       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    492    520       POTENTIAL.
FT   DOMAIN      521    532       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       43    144       SRCR 1.
FT   DOMAIN      153    257       SRCR 2.
FT   DOMAIN      264    366       SRCR 3.
FT   DOMAIN      382    485       SRCR 4.
FT   CARBOHYD     78     78       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   532 AA;  57820 MW;  742533E09576BCB8 CRC64;

  Query Match             6.4%;  Score 520;  DB 1;  Length 532;
  Best Local Similarity   29.1%;  Pred. No. 3.3e-29;
  Matches  152;  Conservative   66;  Mismatches  208;  Indels   96;  Gaps   19;
```

```
Qy      1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLE-----------LR 49
            |||  |   |                      |:::  |  |||  :    |         :|
Db      5 MMLQQYCW-----------------AACLVI---CIAISSVDDVGAEQNYGREAVEGNIR 44

Qy     50 LVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFR---FGQAVTR 106
          . |::|    |:||:   :|| |||  |:   :  | ||||| | :    :|   ||  ||
Db     45 LIHGRTENEGSVEIYHATRWGGVCDWWWHMENANVTCKQLGFPGARQFYRRAYFGAHVT- 103

Qy    107 HGKIWLDDVSCYGNESALWECQHREWG-SHNCYHGEDVGVNCY--GEANLGLRLVDGNNS 163
               |:  ::| |||: | :| || :|     |        || |     |     ||:: |:
Db    104 --TFWVYKMNCLGNETRLEDCYHRPYGRPWLCNAQWAAGVECLPKDEPQGSLRMILGDVP 161

Qy    164 CSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPA------VLRPI 217
          . | :|  :   ||::|    :      | |||:|      | || :             ||
Db    162 NEGTLETFWDGAWGSVCHTDFGTPDGNVACRQMG-----YSRGVKSIKTDGHFGFSTGPI 216

Qy    218 WLDDILCQGNELALWNCRHRGWGN-------HDC--SHNEDVTLTCYDSSDLELRLVGGT 268
           || : |:| |    :   |       |          | |  :|| || : |  : : ::||: |:
Db    217 ILDAVDCEGTEAHITEC------NMPVTPYQHACPYTHNWDVGVVCKPNVEGDIRLMDGS 270

Qy    269 NRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDV----- 323
             ||||:     |||:|   |: | |:|||:| |   |:   :|        | |
Db    271 GPHEGRVEIWHDDAWGTICDDGWDWADANVVCRQAGYRGAVKASGF----KGEDFGFTWA 326

Qy    324 -VWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLAD--------G 374
           :    | |:| |  | ||               | |   | ||:|:   |  : :              |
Db    327 PIHTSFVMCTGVEDRLIDCILRDGWTHSCYHVEDASVVCATDDDDTIEIEPKHTRVRIVG 386

Qy    375 SNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEA------ 428
               |||||  :    |   :|| :|  :  :|  ||    |        :|: ||  | |   .
Db    387 MGQGQGRVEVSLGNGWGRVCDPDWSDHEAKTVCYHAGYK---WGASRAAGSAEVSAPFDL 443

Qy    429 -RDIWINSISCTGNES-ALWDCTYDGKAKRTCFRRSDAGVIC 468
                 |:  |:|:| |:  |   |        |  ||     |  ||:|
Db    444 EAPFIIDGITCSGVENETLSQCQMKVSADMTC-ATGDVGVVC 484
```

RESULT 14

MRCO_MESAU

```
ID   MRCO_MESAU     STANDARD;      PRT;   483 AA.
AC   Q9WUB9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Macrophage receptor MARCO (Macrophage receptor with collagenous
DE   structure).
GN   MARCO.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99242649; PubMed=10224290;
RA   Palecanda A., Paulauskis J., Al-Mutairi E., Imrich A., Qin G.,
RA   Suzuki H., Kodama T., Tryggvason K., Koziel H., Kobzik L.;
RT   "Role of the scavenger receptor MARCO in alveolar macrophage binding
RT   of unopsonized environmental particles.";
RL   J. Exp. Med. 189:1497-1506(1999).
CC   -!- FUNCTION: Bind Gram-positive and Gram-negative bacteria.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF125191; AAD20360.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 1.
DR   Pfam; PF01391; Collagen; 4.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD000007; Collagen; 1.
DR   SMART; SM00202; SR; 1.
DR   PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR   PROSITE; PS50287; SRCR_2; 1.
KW   Collagen; Transmembrane; Receptor; Glycoprotein; Signal-anchor.
FT   DOMAIN        1     48       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     49     69       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       70    483       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      148    383       COLLAGEN-LIKE.
FT   DOMAIN      389    483       SRCR.
FT   DISULFID    412    472       BY SIMILARITY.
FT   DISULFID    425    482       BY SIMILARITY.
FT   DISULFID    452    462       BY SIMILARITY.
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   483 AA;  49621 MW;  C38F18C46505FB1E CRC64;

  Query Match             3.7%;  Score 298;  DB 1;  Length 483;
  Best Local Similarity   52.0%;  Pred. No. 1.1e-13;
  Matches   53;  Conservative   13;  Mismatches   30;  Indels    6;  Gaps    2;

Qy   1034 KRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGE 1093
            | :|:| | :|   || ||:::  ||||||| || :|| | |: ||    :::      ||
Db    387 KEVRIVGGTNR--GRAEIFYNNAWGTICDDNWDNNDATVFCRMLG----YSSGKGFTFGG 440

Qy   1094 GSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVIC 1135
            ||| |||||:|| |||  || |      || |:| | ||||| |
Db    441 GSGNIWLDDVNCQGTEDSLWNCRKNNWGSHNCNHNEDAGVEC 482
```

RESULT 15

MTCO_MOUSE

```
ID   MTCO_MOUSE     STANDARD;      PRT;   518 AA.
AC   Q60754;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Macrophage receptor MARCO (Macrophage receptor with collagenous
DE   structure).
GN   MARCO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95171455; PubMed=7867067;
RA   Elomaa O., Kangas M., Sahlberg C., Tuukkanen J., Sormunen R.,
RA   Liakka A., Thesleff I., Kraal G., Tryggvason K.;
RT   "Cloning of a novel bacteria-binding receptor structurally related to
RT   scavenger receptors and expressed in a subset of macrophages.";
RL   Cell 80:603-609(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99265975; PubMed=10331948;
RA   Kangas M., Brannstrom A., Elomaa O., Matsuda Y., Eddy R., Shows T.B.,
RA   Tryggvason K.;
RT   "Structure and chromosomal localization of the human and murine genes
RT   for the macrophage MARCO receptor.";
RL   Genomics 58:82-89(1999).
CC   -!- FUNCTION: Bind Gram-positive and Gram-negative bacteria.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Expressed in subpopulations of macrophages in
CC       the spleen and the medullary cord of lymph nodes.
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
```

```
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U18424; AAA68638.1; -.
DR   EMBL; AF128423; AAD51136.1; -.
DR   EMBL; AF127927; AAD51136.1; JOINED.
DR   EMBL; AF127928; AAD51136.1; JOINED.
DR   EMBL; AF128169; AAD51136.1; JOINED.
DR   EMBL; AF128170; AAD51136.1; JOINED.
DR   EMBL; AF128171; AAD51136.1; JOINED.
DR   EMBL; AF127601; AAD51136.1; JOINED.
DR   EMBL; AF127602; AAD51136.1; JOINED.
DR   EMBL; AF128419; AAD51136.1; JOINED.
DR   EMBL; AF128420; AAD51136.1; JOINED.
DR   EMBL; AF128421; AAD51136.1; JOINED.
DR   EMBL; AF128422; AAD51136.1; JOINED.
DR   MGD; MGI:1309998; Marco.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF00530; SRCR; 1.
DR   Pfam; PF01391; Collagen; 5.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   ProDom; PD000007; Collagen; 1.
DR   SMART; SM00202; SR; 1.
DR   PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR   PROSITE; PS50287; SRCR_2; 1.
KW   Collagen; Transmembrane; Receptor; Glycoprotein; Signal-anchor.
FT   DOMAIN        1     48       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     49     69       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       70    518       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      149    418       COLLAGEN-LIKE.
FT   DOMAIN      423    518       SRCR.
FT   DISULFID    446    507       BY SIMILARITY.
FT   DISULFID    459    517       BY SIMILARITY.
FT   DISULFID    487    497       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   518 AA;  52730 MW;  B09E7601ECA23637 CRC64;

  Query Match              3.6%;  Score 290.5;  DB 1;  Length 518;
  Best Local Similarity   49.5%;  Pred. No. 4.2e-13;
  Matches   51;  Conservative   16;  Mismatches   31;  Indels    5;  Gaps    2;

Qy   1034 KRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGE 1093
          :|:|:: | :|  || |:|::  ||||||| || :|| | |: ||          : :|
Db    421 QRVRIMGGTNR--GRAEVYYNNEWGTICDDDWDNNDATVFCRMLGYS---RGRALSSYGG 475

Qy   1094 GSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICS 1136
          ||| ||||::|| |||: || |    || |:| | ||||| ||
Db    476 GSGNIWLDNVNCRGTENSLWDCSKNSWGNHNCVHNEDAGVECS 518

Search completed: May 12, 2003, 06:37:50
Job time : 28 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:36:46 ; Search time 24 Seconds
(without alignments)
1781.313 Million cell updates/sec

Title: US-09-759-130B-381
Perfect score: 8138
Sequence: 1 MMLPQNSWHIDFGRCCCHQN...........CEDASDTSLLGVLPASEATK 1453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2435	29.9	1785	4	US-09-341-587-3	Sequence 3, Appli
2	1048.5	12.9	1290	1	US-08-470-350B-2	Sequence 2, Appli
3	718.5	8.8	666	4	US-09-341-587-1	Sequence 1, Appli
4	691.5	8.5	346	3	US-09-034-916-2	Sequence 2, Appli
5	629	7.7	754	4	US-09-276-400-8	Sequence 8, Appli
6	629	7.7	754	4	US-09-448-076-8	Sequence 8, Appli
7	629	7.7	754	4	US-09-702-572-8	Sequence 8, Appli
8	622.5	7.6	753	4	US-09-276-400-2	Sequence 2, Appli
9	622.5	7.6	753	4	US-09-448-076-2	Sequence 2, Appli
10	622.5	7.6	753	4	US-09-702-572-2	Sequence 2, Appli
11	607.5	7.5	774	4	US-09-276-400-7	Sequence 7, Appli
12	607.5	7.5	774	4	US-09-448-076-7	Sequence 7, Appli
13	607.5	7.5	774	4	US-09-702-572-7	Sequence 7, Appli
14	462	5.7	100	3	US-09-034-916-15	Sequence 15, Appl
15	432	5.3	100	3	US-09-034-916-17	Sequence 17, Appl
16	412	5.1	100	3	US-09-034-916-13	Sequence 13, Appl
17	347	4.3	100	3	US-09-034-916-11	Sequence 11, Appl
18	346	4.3	102	3	US-09-034-916-16	Sequence 16, Appl
19	338	4.2	100	3	US-09-034-916-10	Sequence 10, Appl
20	337.5	4.1	101	3	US-09-034-916-9	Sequence 9, Appli
21	326.5	4.0	95	3	US-09-034-916-7	Sequence 7, Appli
22	311.5	3.8	585	1	US-08-477-674-10	Sequence 10, Appl
23	311.5	3.8	585	1	US-08-473-791-10	Sequence 10, Appl
24	311.5	3.8	585	2	US-08-316-714-10	Sequence 10, Appl
25	311.5	3.8	585	3	US-08-473-673-10	Sequence 10, Appl
26	311	3.8	100	3	US-09-034-916-14	Sequence 14, Appl
27	296	3.6	102	3	US-09-034-916-5	Sequence 5, Appli
28	290.5	3.6	489	2	US-08-794-795-7	Sequence 7, Appli
29	290.5	3.6	489	4	US-09-249-200-7	Sequence 7, Appli
30	290.5	3.6	518	1	US-08-392-367B-2	Sequence 2, Appli
31	290.5	3.6	518	3	US-08-893-467A-2	Sequence 2, Appli
32	290	3.6	451	1	US-08-154-365-2	Sequence 2, Appli
33	288.5	3.5	495	2	US-08-794-795-2	Sequence 2, Appli
34	288.5	3.5	495	4	US-09-249-200-2	Sequence 2, Appli
35	288.5	3.5	520	2	US-08-794-795-6	Sequence 6, Appli
36	288.5	3.5	520	4	US-09-249-200-6	Sequence 6, Appli
37	288	3.5	451	1	US-08-453-117-2	Sequence 2, Appli
38	288	3.5	451	2	US-08-948-222-2	Sequence 2, Appli
39	288	3.5	451	2	US-08-973-145-2	Sequence 2, Appli
40	288	3.5	451	4	US-09-276-400-10	Sequence 10, Appl
41	288	3.5	451	4	US-09-448-076-10	Sequence 10, Appl
42	288	3.5	451	4	US-09-702-572-10	Sequence 10, Appl
43	288	3.5	451	5	PCT-US96-08081-2	Sequence 2, Appli
44	278	3.4	101	4	US-09-518-046-16	Sequence 16, Appl
45	273	3.4	101	4	US-09-518-046-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-341-587-3
; Sequence 3, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-3

```
  Query Match             29.9%;  Score 2435;  DB 4;  Length 1785;
  Best Local Similarity   36.1%;  Pred. No. 3.9e-206;
  Matches  533;  Conservative  155;  Mismatches  415;  Indels  374;  Gaps   28;

Qy       43 GTD--LELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAM--- 97
            |:|  | |||||||| | | ||: ::| ||||||| |:|  : |||:|||| :: :
Db       95 GSDSGLALRLVNGDGRCQGRVEILYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGN 154

Qy       98 FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCY--------- 148
              |||     | | |||| | |:|| || | |  | |||| |||| || |
Db      155 AWFGQG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAQPQSTLR 211

Qy      149 ----------------GEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVV 192
                             |::| |||:|  : | ||||| ::  |||:||| |: | | ||
Db      212 PESWPVRISPPVPTEGSESSLALRLVNGGDRCRGRVEVLYRGSWGTVCDDYWDTNDANVV 271

Qy      193 CRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTL 252
            ||||||  :  : |         || |||: | |:|. ||:| | ||  |:| |:||  :
Db      272 CRQLGCGWAMSAPGNAQFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLTHNCGHSEDAGV 331

Qy      253 TCY---------------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHK 290
             |                         | | |||| | :|| ||||:  :| |||||
Db      332 ICSAPQSRPTPSPDTWPTSHASTAGPESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDS 391

Qy      291 WNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFD 350
            |: : |:|||:||||| |    |      ||  : || | ||| ||:|| | |:| :: :
Db      392 WDTSDANVVCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGYESYLWSCPHNGWLSHN 451

Qy      351 CLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQL 410
            | |  |  ||||                           | |
```

```
Db    452 CQHSEDAGVICSAA----------------------HSWST------------------470

Qy    411 GCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470
                        || :           |  |:||:
Db    471 -------------PSPDTLPTITLPASTVGSESS------------------------- 491

Qy    471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTY 530
             | ||||       | ||:|| ||| ||||| | | | :| |||:| |||   |     |
Db    492 ---LALRLVNGGDRCQGRVEVLYQGSWGTVCDDSWDTNDANVVCRQPGCGWAMSAPGNAR 548

Qy    531 FKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDA--------TW- 581
          | : |||| |||| | |:||  | | |:||  ||| || |  | ||            ||
Db    549 FGQGSGPIVLDDVRCSGHESYPWSCPHNGWLSHNCGHSEDAGVICSASQSRPTPSPDTWP 608

Qy    582 -------------GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPS 628
                       |||| | :|| ||:|| ::| |||||| |::  | ||| || |
Db    609 TSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANVVCRQLGCGW 668

Qy    629 SIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICS----- 683
          ::    |      | | | |||| | | || |||| ::| :::| | || |||||
Db    669 AMSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQ 728

Qy    684 -----------------DASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEV 726
                           | : |||| |  || |:||| :|: | :| : |  | | |
Db    729 PTPSPDTWPTSHASTAGSESSLALRLVNGGDRCQGRVEVLYRGSWGTVCDDYWDTNDANV 788

Qy    727 VCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEAS 786
          ||||| || |     | : :  |:: :  |:| |: || |       |     | :|
Db    789 VCRQLGCGWATSAPGNARFGQGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAG 848

Qy    787 LICSAHR-QPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLS 845
          :|||| : ||       .              ||| :
Db    849 VICSASQSQP--------------TPSPDTWPT-------------------------- 867

Qy    846 VGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLVNG 905
                            :     |||: |||                         |||||
Db    868 --------------SRASTAGSESTLAL--------------------------RLVNG 886

Qy    906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGH 965
            :| |:||:   | ||::|| :||   || |:|||| || |:|   |       |: |   :
Db    887 GDRCRGRVEVLYQGSWGTVCDDYWDTNDANVVCRQLGCGWAMSAPGNAQFGQGSGPIVLD 946

Qy    966 RFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPYLSAVP 1023
            | |:|| | :|        |  | |      |||: :| :|  |          :| |:
Db    947 DVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQSTPRPDTWLTTNLPALTVGS 1006

Qy   1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAF 1083
            | |           ||||:| || ||||: : | |||:||| || :||:|||::|||| |
Db   1007 ESSL-------ALRLVNGGDRCRGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAM 1059

Qy   1084 NATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRL 1143
           :|  :| ||:||||| |||: |:| ||:|| || :|| |:|| || | ||||||| | ::
Db   1060 SAPGNARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLTHNCGHHEDAGVICS---ATQI 1116

Qy   1144 YSET-------------ETESCAGRLEVFY-NGTWGS---------------------- 1166
           | |               : :| | |   || :||:  |
Db   1117 NSTTTDWWHPTTTTTARPSSNCGGFL--FYASGTFSSPSYPAYYPNNAKCVWEIEVNSGY 1174

Qy   1167 ---VGRRNI--------------------TTAIAGIVC---RQLGCGENGVVSL---AP 1196
             :|  |:                      :: : | :|  ||:     ::: :    :
Db   1175 RINLGFSNLKLEAHHNCSFDYVEIFDGSLNSSLLLGKICNDTRQIFTSSYNRMTIHFRSD 1234

Qy   1197 LSKTGSGFM-WVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTEC 1255
           :|   :||: | :                          | |:: |       : :   |
Db   1235 ISFQNTGFLAWYN------------------------SFPSDATL----RLVNLNSSYGLC 1267

Qy   1256 SGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRC 1315
           :|||||:| |:|||||||||  : |||||:|||| |::|| :| || |:| |.| |||: |
Db   1268 AGRVEIYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVEC 1327

Qy   1316 KGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352
            | || || |   :  |   :| |:||||| ||| |  :
Db   1328 SGTESTLWQCRNRGWFSHNCNHREDAGVICSGNHLST 1364
```

RESULT 2
US-08-470-350B-2
; Sequence 2, Application US/08470350B
; Patent No. 5684126
; GENERAL INFORMATION:
; APPLICANT: Li, Xiao
; APPLICANT: Snyder, Solomon H
; TITLE OF INVENTION: Ebnerin: A Secreted von Ebner's Gland
; TITLE OF INVENTION: Protein Associated with Taste Buds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,350B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wolffe, Susan A
; REGISTRATION NUMBER: 33,568
; REFERENCE/DOCKET NUMBER: 01107.48790
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1290 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-470-350B-2

```
  Query Match             12.9%;  Score 1048.5;  DB 1;  Length 1290;
  Best Local Similarity   29.7%;  Pred. No. 1.1e-83;
  Matches  289;  Conservative  114;  Mismatches  290;  Indels  281;  Gaps   34;

Qy    286 VCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSG 345
          :|   |:   |:|||:||||| ||    |      |     : || |:| |:|::|| | | |
Db      1 MCDDSWDINDANVVCRQLGCGWALSAPGSAQFGQGLGPIVLDDVACRGHEAYLWSCSHRG 60

Qy    346 TVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALV 405
          :: :| || |  |||||            :  |:  |         ||    :    |  :
Db     61 WLSHNCGHQEDAGVICSD---------SQTSSPTPG---------WW---NPGGTNNDVIY 100

Qy    406 VCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAG 465
           ::    :: |:   | |          :    : |   ||    :       :|:|
Db    101 DTQET--------TETSQTSSPTPDWW-------NHGGTINDVIYD---TQETTEGTDSG 142

Qy    466 VICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHV 525
                  | :|||      | ||:|: ||| ||||| | |     :| |||:||||| :
Db    143 --------LAVRLVNGGDRCRGRVEILYQGSWGTVCDDSWDINDANVVCRQLGCGWALSA 194

Qy    526 FGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCS--------- 576
           |    | : || | ||||:| |:|: :| | | || ||| |:|| | ||
Db    195 PGSAQFGQGSGSIVLDDVACRGHEAYLWSCSHRGWLSHNCGHQEDAGVICSYSQTSSPTP 254

Qy    577 ------------------GDATWG-------------LRLVGGSNRCSGRLEVYFQGRWG 605
                             | ::|             :||| | :|| ||:|: :|| ||
Db    255 DSQTSSPTPGWWNPGGTNNDVSYGPEQTTDATDSGLAVRLVNGGDRCQGRVEILYQGSWG 314

Qy    606 TVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSCRNS 665
          ||||| |::| | ||| || |   ::    |  :   |   | | |||:| | |: |||| :
```

```
Db    315 TVCDDSWDTKDANVVCRQLVCGWALSAPGSAHFGQGSGSIVLDDVACTGHEAYLWSCSHR 374

Qy    666 GWGNNDCSHSEDVGVICSDAS-------DM------ELRLVGGSSRCAGKVEVNVQGAVG 712
          || :::| | || |||||||        ||      |      ::: :  |       :
Db    375 GWLSHNCGHHEDAGVICSDAQTQSTTWPDMWPTTTPETTTDWWTTKYSSSVPTTQFPTI- 433

Qy    713 ILCANGWGMNIAEVVCRQLECGSAIRVS----REPHFTERTLHILMSNSGCTGGEASLWD 768
             |: |       |  .      ||  :  :       |::          :|: |      ||
Db    434 ---ADWWTTPSPEYT-----CGGLLTLPYGQFSSPYYPGS----YPNNARC------LW- 474

Qy    769 CIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSL-- 826
                            :  |:  :  :|  |:   |                 |: |: |
Db    475 -----------------KIFVSSMNRVTVVFTDVQLEGG------------CNYDYILVF 505

Qy    827 ----HAANVLCRELNCGDAISLSVGDHFGKG--------------NGLTWAEKFQCE--- 865
               : :::: |                  | | |                 :|        || :
Db    506 DGPENNSSLIAR-----------VCDGFNGSFTSTQNFMSVVFITDGSVTRRGFQADYYS 554

Qy    866 ----GSETHLALCPIV------QHPEDTC--IHSREVGVVCSRYTDVRLVNGKSQCDGQV 913
                || |  |||           || ||  : :   |   | |
Db    555 TPISTSTTSPTTFPIVTDWWTTPSPEYTCGGLLTLPYGQFSSPY---------------- 598

Qy    914 EINVLGHWGSLCDTHWDPEDARVLCR---------------QLSCGT----ALSTTGGK 953
                 : ||        | :|| | :                  ||   |        |    | :
Db    599 ------YPGSY------PNNARCLWKIFVPSMNRVTVVFTDVQLEGGCNYDYILGFDGPE 646

Qy    954 YIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSV--ICTGSLTQPLF--- 1008
          |       ||         | |:       : |            | :||   |    ||:|:  |
Db    647 YNSSLIARV------CDGSNGSFTSTQ-----------NFMSVVFITDGSVTRRGFQAD 688

Qy   1009 ----PCLANVSDPYLSAVPEG--SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICD 1062
               |   : : |      :  |   |:|:      ||||:| :|| ||||| : | |       |
Db    689 YYSTPIRTSTTPPTTFPIITGNDSSLV------LRLVNGTNRCEGRVEILYRGSWVPCAD 742

Qy   1063 DGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQ 1122
           | ||::||:|||::|||| | :|   :| ||:||| | |||::|:| ||||| |     ||
Db    743 DSWDINDANVVCRQLGCGSALSAPGNAWFGQGSGLIVLDDVSCSGYESHLWNCRHPGWLV 802

Qy   1123 HDCRHKEDAGVICS 1136
          |:||| ||||||||
Db    803 HNCRHVEDAGVICS 816
```

RESULT 3
US-09-341-587-1
; Sequence 1, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 666
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-341-587-1

```
  Query Match             8.8%;  Score 718.5;  DB 4;  Length 666;
  Best Local Similarity   35.8%;  Pred. No. 6.7e-55;
  Matches  162;  Conservative   63;  Mismatches  118;  Indels  109;  Gaps   13;

Qy    971 GNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPYLSAVPEGSAL 1028
          |:|| | :|        :  | |      |||: : :|  |            : | |:   | |
Db      1 GHESYLWSCPHNGWLSHNCGHHEDAGVICSAAQSQSTPRPDWTLTTNLPALTVGSESSL- 59

Qy   1029 ICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVS 1088
                 ||||:| || ||||: : | |||:||| || :||:|||::|||| | :|  : .
Db     60 ------ALRLVNGGDRCRGRVEVLYRGSWGTVCDDSWDTNDANVVCRQLGCGWAMSAPGN 113

Qy   1089 AHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSET- 1147
           | ||:||||| |||: |:| ||:|| || :|| |:| | ||||||||   | :: | |
Db    114 ARFGQGSGPIVLDDVRCSGNESYLWSCPHKGWLTHNCGHHEDAGVICS---ATQINSTTT 170

Qy   1148 ------------ETESCAGRLEVFY-NGTWGS-------------------------VG 1168
                      : :| |  |   || :||: |                           :|
Db    171 DWWHPTTTTARPSSNCGGFL--FYASGTFSSPSYPAYYPNNAKCVWEIEVNSGYRINLG 228

Qy   1169 RRNI--------------------TTAIAGIVC---RQLGCGENGVVSL---APLSKTG 1201
           |:                           :: :  | :|      ||:         :::   : :|
Db    229 FSNLKLEAHHNCSFDYVEIFDGSLNSSLLGKICNDTRQIFTSSYNRMTIHFRSDISFQN 288

Qy   1202 SGFM-WVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260
          :||: | :                       | |:: |          : :        |:||||
Db    289 TGFLAWYN----------------------SFPSDATL----RLVNLNSSYGLCAGRVE 321

Qy   1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
          |:| |:||||||||| : ||||||:||||| |::|| :| || |:| | |||: | | | ||
Db    322 IYHGGTWGTVCDDSWTIQEAEVVCRQLGCGRAVSALGNAYFGSGSGPITLDDVECSGTES 381

Qy   1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKS 1352
           || |  : |   :| |:||||| ||| | :
Db    382 TLWQCRNRGWFSHNCNHREDAGVICSGNHLST 413
```

RESULT 4
US-09-034-916-2
; Sequence 2, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
;   APPLICANT: GEBE, JOHN A.
;   APPLICANT: SIADAK, ANTHONY W.
;   APPLICANT: ARUFFO, ALEJANDRO A.
;   TITLE OF INVENTION: SpALPHA: A NOVEL SCAVENGER RECEPTOR
;   TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
;   TITLE OF INVENTION: ANTIBODIES THERETO
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ROBINS & ASSOCIATES
;   STREET: 90 MIDDLEFIELD ROAD, SUITE 200
;   CITY: MENLO PARK
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/034,916
;   FILING DATE: 04-MAR-1998
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/039,956
;   FILING DATE: 06-MAR-1997
;   ATTORNEY/AGENT INFORMATION:
;   NAME: ROBINS, ROBERTA L.
;   REGISTRATION NUMBER: 33,208
;   REFERENCE/DOCKET NUMBER: 5998-0020
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 325-7812
;   TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 346 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-034-916-2

```
  Query Match             8.5%;  Score 691.5;  DB 3;  Length 346;
  Best Local Similarity   41.3%;  Pred. No. 6.1e-53;
  Matches  148;  Conservative   40;  Mismatches  139;  Indels   31;  Gaps    9;

Qy         21 LFSAVVTCILLLNSC----FLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDG 76
Db          3 LFS------LILAICTRPGFLASPSG-----VRLVGGLHRCEGRVEVEQKGQWGTVCDDG 51

Qy         77 WNTTASTVVCKQLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWG 133
Db         52 WDIKDVAVLCRELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEE-- 109

Qy        134 SHNCYHGEDVGVNC------YGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLN 187
Db        110 VYDCSHDEDAGASCENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLR 169

Qy        188 TAAVVCRQLGCPSSFISSGVVNSPAVLR-PIWLDDILCQGNELALWNCRHRGWGNHDCSH 246
Db        170 AAKVVCRQLGCGRAVLTQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNH 229

Qy        247 NEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCG 306
Db        230 DEDTWVECEDPFD--LRLVGGDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCG 287

Qy        307 TAL--HFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS 362
Db        288 KSLSPSFRDRKCYGPGVGRIWLDNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICS 345
```

RESULT 5
US-09-276-400-8
; Sequence 8, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 754
;   TYPE: PRT
;   ORGANISM: murine lysyl oxidase-related protein
US-09-276-400-8

```
  Query Match             7.7%;  Score 629;  DB 4;  Length 754;
  Best Local Similarity   31.0%;  Pred. No. 6.7e-47;
  Matches  155;  Conservative   72;  Mismatches  183;  Indels   90;  Gaps   16;

Qy         56 PCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKI 110
Db         54 PYEGRVEIQRAGEWGTICDDDFTLQAAHVLCRELG--FTEATGWTHSAKYGPGT---GRI 108

Qy        111 WLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNC-------YGEANL---------- 153
Db        109 WLDNLSCRGTEGSVTECASRGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHQLQVE 168

Qy        154 ----------GLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS-- 201
Db        169 EVRLRPAVEWGRRPLP---VTEGLVEVRLPEGWSQVCDKGWSAHNSHVVCGMLGFPGEKR 225

Qy        202 -------FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHD---CSHNEDVT 251
Db        226 VNMAFYRMLAQKKQHS------FGLHSVACVGTEAHLSLCSLEFYRANDTTRCSGGNPAV 279

Qy        252 LTCY-----------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNN 293
Db        280 VSCVLGPLYATFTGQKKQQHSKPQGEARVRLKGGAHQGEGRVEVLKAGTWGTVCDRKWDL 339

Qy        294 AAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLH 353
Db        340 QAASVVCPELGFGTAREALSGARMGQGMGAIHLSEVRCSGQEPSLWRCPSKNITAEDCSH 399

Qy        354 QNDVSVICS---DGADLELRLADGSNNCSGRVEVRI----HEQWWTICDQNWKNEQALVV 406
Db        400 SQDAGVRCNLPYTGVETKIRLSGGRSRYEGRVEVQIGIPGHLRWGLICGDDWGTLEAMVA 459

Qy        407 CKQLGCPFSVFGSRRA--KPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR---R 461
Db        460 CRQLGLGYANHGLQETWYWDSGNVTEVVMSGVRCTGSELSLNQCAHH-SSHITCKKTGTR 518

Qy        462 SDAGVICSDKADLDLRLVGA 481
Db        519 FTAGVICSETAS-DLLLHSA 537
```

RESULT 6
US-09-448-076-8
; Sequence 8, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: MNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 754
;   TYPE: PRT
;   ORGANISM: murine lysyl oxidase-related protein
US-09-448-076-8

```
  Query Match             7.7%;  Score 629;  DB 4;  Length 754;
  Best Local Similarity   31.0%;  Pred. No. 6.7e-47;
  Matches  155;  Conservative   72;  Mismatches  183;  Indels   90;  Gaps   16;

Qy         56 PCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKI 110
Db         54 PYEGRVEIQRAGEWGTICDDDFTLQAAHVLCRELG--FTEATGWTHSAKYGPGT---GRI 108

Qy        111 WLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNC-------YGEANL---------- 153
Db        109 WLDNLSCRGTEGSVTECASRGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHQLQVE 168

Qy        154 ----------GLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS-- 201
Db        169 EVRLRPAVEWGRRPLP---VTEGLVEVRLPEGWSQVCDKGWSAHNSHVVCGMLGFPGEKR 225

Qy        202 -------FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHD---CSHNEDVT 251
Db        226 VNMAFYRMLAQKKQHS------FGLHSVACVGTEAHLSLCSLEFYRANDTTRCSGGNPAV 279

Qy        252 LTCY-----------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNN 293
Db        280 VSCVLGPLYATFTGQKKQQHSKPQGEARVRLKGGAHQGEGRVEVLKAGTWGTVCDRKWDL 339

Qy        294 AAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLH 353
Db        340 QAASVVCPELGFGTAREALSGARMGQGMGAIHLSEVRCSGQEPSLWRCPSKNITAEDCSH 399

Qy        354 QNDVSVICS---DGADLELRLADGSNNCSGRVEVRI----HEQWWTICDQNWKNEQALVV 406
Db        400 SQDAGVRCNLPYTGVETKIRLSGGRSRYEGRVEVQIGIPGHLRWGLICGDDWGTLEAMVA 459

Qy        407 CKQLGCPFSVFGSRRA--KPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR---R 461
Db        460 CRQLGLGYANHGLQETWYWDSGNVTEVVMSGVRCTGSELSLNQCAHH-SSHITCKKTGTR 518
```

```
Qy    462 SDAGVICSDKADLDLRLVGA 481
            ||||||: |  || |  |
Db    519 FTAGVICSETAS-DLLLHSA 537


RESULT 7
US-09-702-572-8
; Sequence 8, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
;  APPLICANT:  Khodadoust, Mehran
;  TITLE OF INVENTION:  NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE:  MNI-073
;  CURRENT APPLICATION NUMBER: US/09/702,572
;  CURRENT FILING DATE:  2000-10-31
;  PRIOR APPLICATION NUMBER:  09/276,400
;  PRIOR FILING DATE:  1999-03-25
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 754
;   TYPE: PRT
;   ORGANISM: murine lysyl oxidase-related protein
US-09-702-572-8

  Query Match             7.7%;  Score 629;  DB 4;  Length 754;
  Best Local Similarity   31.0%;  Pred. No. 6.7e-47;
  Matches  155;  Conservative   72;  Mismatches  183;  Indels   90;  Gaps   16;

Qy     56 PCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKI 110
           |  | ||::  |:|||:|||  :    |: |:|::||  |: |        ::|        |:|
Db     54 PYEGRVEIQRAGEWGTICDDDFTLQAAHVLCRELG--FTEATGWTHSAKYGPGT---GRI 108

Qy    111 WLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNC-------YGEANL---------- 153
           |||::||  | |  :: ||   | ||: :| |  || || |             : ::|:
Db    109 WLDNLSCRGTEGSVTECASRGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHQLQVE 168

Qy    154 ---------GLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS-- 201
                     | | :        | |||:   | |   :|| ||:`: :  |||   || |
Db    169 EVRLRPAVEWGRRPLP---VTEGLVEVRLPEGWSQVCDKGWSAHNSHVVCGMLGFPGEKR 225

Qy    202 -------FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHD---CSHNEDVT 251
                  ::     :|          |  :  |  |  |   |   |      :  :|    ||
Db    226 VNMAFYRMLAQKKQHS------FGLHSVACVGTEAHLSLCSLEFYRANDTTRCSGGNPAV 279

Qy    252 LTCY-----------------DSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNN 293
           ::|                         :  :|| || ::  ||||:    | |||||  ||:
Db    280 VSCVLGPLYATFTGQKKQQHSKPQGEARVRLKGGAHQGEGRVEVLKAGTWGTVCDRKWDL 339

Qy    294 AAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLH 353
            || ||| :|| |||             :  |     : |  | ||| |  || |           || |
Db    340 QAASVVCPELGFGTAREALSGARMGQGMGAIHLSEVRCSGQEPSLWRCPSKNITAEDCSH 399

Qy    354 QNDVSVICS---DGADLELRLADGSNNCSGRVEVRI----HEQWWTICDQNWKNEQALVV 406
             |  | |:      | : ::||: |  :       |||||:|      | :|  ||  :|   :|:|
Db    400 SQDAGVRCNLPYTGVETKIRLSGGRSRYEGRVEVQIGIPGHLRWGLICGDDWGTLEAMVA 459

Qy    407 CKQLGCPFSVFGSRRA--KPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR---R 461
           |:|||   ::   | :          |       :: :: :  |||:| :|  | :     :  || :    |
Db    460 CRQLGLGYANHGLQETWYWDSGNVTEVVMSGVRCTGSELSLNQCAHH-SSHITCKKTGTR 518

Qy    462 SDAGVICSDKADLDLRLVGA 481
            ||||||: |  || |  |
Db    519 FTAGVICSETAS-DLLLHSA 537


RESULT 8
US-09-276-400-2
; Sequence 2, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
;  APPLICANT: Khodadoust, Mehran
;  TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION: USES THEREFOR
;  FILE REFERENCE: MNI-073
;  CURRENT APPLICATION NUMBER: US/09/276,400
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 753
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-276-400-2

  Query Match             7.6%;  Score 622.5;  DB 4;  Length 753;
  Best Local Similarity   31.2%;  Pred. No. 2.5e-46;
  Matches  167;  Conservative   71;  Mismatches  202;  Indels   95;  Gaps   19;

Qy     30 LLLNSCFLISSFNGT-------------DLELRLVN-GDGPCSGTVEVKFQGQWGTVCD 74
           |||  | | ||  |:                        |  ||        |  | ||::  |:|||:||
Db     14 LLL--CLLCSSCLGSPSPSTGPEKKAGSQGLRFRLAGFPRKPYEGRVEIQRAGEWGTICD 71

Qy     75 DGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECQH 129
           | :    |: ::|::||  |: |            ::|       |:||||::|| | |  :: ||
Db     72 DDFTLQAAHILCRELG--FTEATGWTHSAKYGPGT---GRIWLDNLSCSGTEQSVTECAS 126

Qy    130 REWGSHNCYHGEDVGVNC-------YGEANL-------------------GLRLVDGNN 162
           | ||: :| | || || |          : ::|:                      | |  :
Db    127 RGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHHLQVEEVRIRPAVGWGRRPLP--- 183

Qy    163 SCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLR---PIWL 219
              | |||:  :  |  :|| ||: : : |||  || ||          :                |      |
Db    184 VTEGLVEVRLPDGWSQVCDKGWSAHNSHVVCGMLGFPSEKRVNAAFYRLLAQRQQHSFGL 243

Qy    220 DDILCQGNELALWNCRHRGWGNHD---CSHNEDVTLTC-----YDSS------------ 258
             : |  |  |   |     :  :|    |          :: |      | :|
Db    244 HGVACVGTEAHLSLCSLEFYRANDTARCPGGGPAVVSCVPGPVYAASSGQKKQQQSKPQG 303

Qy    259 DLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
           :  :|| ||  :     ||||:       |||||   ||:   || |||::|| |:|                 :
Db    304 EARVRLKGGAHPGEGRVEVLKASTWGTVCDRKWDLHAASVVCRELGFGSAREALSGARMG 363

Qy    319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS---DGADLELRLADGS 375
             |   :  |  | |||  |   || |  |               || |   |  |  |  |:    ||:  :||: |
Db    364 QGMGAIHLSEVRCSGQELSLWKCPHKNITAEDCSHSQDAGVRCNLPYTGAETRIRLSGGR 423

Qy    376 NNCSGRVEVRIHE----QWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA--KPSNEAR 429
            :    |||||:|      :| |||  :|     :|:|  |:|:|||   ::   | :        |
Db    424 SQHEGRVEVQIGGPGPLRWGLICGDDWGTLEAMVACRQLGLGYANHGLQETWYWDSGNIT 483

Qy    430 DIWINSISCTGNESALWDCTYDGKAKRTCFR---RSDAGVICSDKADLDLRLVGA 481
           :: :: :  ||| | :|  | : |          || |        |  ||||||: |  || |  |
Db    484 EVVMSGVRCTGTELSLDQCAHHG-THITCKRTGTRFTAGVICSETAS-DLLLHSA 536


RESULT 9
US-09-448-076-2
; Sequence 2, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
;  APPLICANT: Khodadoust, Mehran et al.
;  TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
;  FILE REFERENCE: MNI-073CP
;  CURRENT APPLICATION NUMBER: US/09/448,076
;  CURRENT FILING DATE: 1999-11-23
;  EARLIER APPLICATION NUMBER: 60/117,580
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 09/276,400
;  EARLIER FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 753
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-448-076-2

  Query Match              7.6%;  Score 622.5;  DB 4;  Length 753;
  Best Local Similarity   31.2%;  Pred. No. 2.5e-46;
  Matches  167;  Conservative   71;  Mismatches  202;  Indels   95;  Gaps   19;

Qy      30 LLLNSCFLISSFNGT--------------DLELRLVN-GDGPCSGTVEVKFQGQWGTVCD 74
Db      14 LLL--CLLCSSCLGSPSPSTGPEKKAGSQGLRFRLAGFPRKPYEGRVEIQRAGEWGTICD 71

Qy      75 DGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECQH 129
Db      72 DDFTLQAAHILCRELG--FTEATGWTHSAKYGPGT---GRIWLDNLSCSGTEQSVTECAS 126

Qy     130 REWGSHNCYHGEDVGVNC-------YGEANL-------------------GLRLVDGNN 162
Db     127 RGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHHLQVEEVRIRPAVGWGRRPLP--- 183

Qy     163 SCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLR---PIWL 219
Db     184 VTEGLVEVRLPDGWSQVCDKGWSAHNSHVVCGMLGFPSEKRVNAAFYRLLAQRQQHSFGL 243

Qy     220 DDILCQGNELALWNCRHRGWGNHD---CSHNEDVTLTC-----YDSS------------ 258
Db     244 HGVACVGTEAHLSLCSLEFYRANDTARCPGGGPAVVSCVPGPVYAASSGQKKQQQSKPQG 303

Qy     259 DLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
Db     304 EARVRLKGGAHPGEGRVEVLKASTWGTVCDRKWDLHAASVVCRELGFGSAREALSGARMG 363

Qy     319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS---DGADLELRLADGS 375
Db     364 QGMGAIHLSEVRCSGQELSLWKCPHKNITAEDCSHSQDAGVRCNLPYTGAETRIRLSGGR 423

Qy     376 NNCSGRVEVRIHE----QWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA--KPSNEAR 429
Db     424 SQHEGRVEVQIGGPGPLRWGLICGDDWGTLEAMVACRQLGLGYANHGLQETWYWDSGNIT 483

Qy     430 DIWINSISCTGNESALWDCTYDGKAKRTCFR---RSDAGVICSDKADLDLRLVGA 481
Db     484 EVVMSGVRCTGTELSLDQCAHHG-THITCKRTGTRFTAGVICSETAS-DLLLHSA 536


RESULT 10
US-09-702-572-2
; Sequence 2, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
;  APPLICANT:  Khodadoust, Mehran
;  TITLE OF INVENTION:  NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE:  MNI-073
;  CURRENT APPLICATION NUMBER: US/09/702,572
;  CURRENT FILING DATE:  2000-10-31
;  PRIOR APPLICATION NUMBER:  09/276,400
;  PRIOR FILING DATE:  1999-03-25
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 753
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-702-572-2

  Query Match              7.6%;  Score 622.5;  DB 4;  Length 753;
  Best Local Similarity   31.2%;  Pred. No. 2.5e-46;
  Matches  167;  Conservative   71;  Mismatches  202;  Indels   95;  Gaps   19;

Qy      30 LLLNSCFLISSFNGT--------------DLELRLVN-GDGPCSGTVEVKFQGQWGTVCD 74
Db      14 LLL--CLLCSSCLGSPSPSTGPEKKAGSQGLRFRLAGFPRKPYEGRVEIQRAGEWGTICD 71

Qy      75 DGWNTTASTVVCKQLGCPFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECQH 129
Db      72 DDFTLQAAHILCRELG--FTEATGWTHSAKYGPGT---GRIWLDNLSCSGTEQSVTECAS 126

Qy     130 REWGSHNCYHGEDVGVNC-------YGEANL-------------------GLRLVDGNN 162
Db     127 RGWGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHHLQVEEVRIRPAVGWGRRPLP--- 183

Qy     163 SCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLR---PIWL 219
Db     184 VTEGLVEVRLPDGWSQVCDKGWSAHNSHVVCGMLGFPSEKRVNAAFYRLLAQRQQHSFGL 243

Qy     220 DDILCQGNELALWNCRHRGWGNHD---CSHNEDVTLTC-----YDSS------------ 258
Db     244 HGVACVGTEAHLSLCSLEFYRANDTARCPGGGPAVVSCVPGPVYAASSGQKKQQQSKPQG 303

Qy     259 DLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
Db     304 EARVRLKGGAHPGEGRVEVLKASTWGTVCDRKWDLHAASVVCRELGFGSAREALSGARMG 363

Qy     319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS---DGADLELRLADGS 375
Db     364 QGMGAIHLSEVRCSGQELSLWKCPHKNITAEDCSHSQDAGVRCNLPYTGAETRIRLSGGR 423

Qy     376 NNCSGRVEVRIHE----QWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRA--KPSNEAR 429
Db     424 SQHEGRVEVQIGGPGPLRWGLICGDDWGTLEAMVACRQLGLGYANHGLQETWYWDSGNIT 483

Qy     430 DIWINSISCTGNESALWDCTYDGKAKRTCFR---RSDAGVICSDKADLDLRLVGA 481
Db     484 EVVMSGVRCTGTELSLDQCAHHG-THITCKRTGTRFTAGVICSETAS-DLLLHSA 536


RESULT 11
US-09-276-400-7
; Sequence 7, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
;  APPLICANT: Khodadoust, Mehran
;  TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION: USES THEREFOR
;  FILE REFERENCE: MNI-073
;  CURRENT APPLICATION NUMBER: US/09/276,400
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 774
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-276-400-7

  Query Match              7.5%;  Score 607.5;  DB 4;  Length 774;
  Best Local Similarity   27.9%;  Pred. No. 5.6e-45;
  Matches  196;  Conservative   86;  Mismatches  258;  Indels  163;  Gaps   28;

Qy     260 LELRLVGGTNR-CMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
Db      56 IQLRLAGQKRKHSEGRVEVYYDGQWGTVCDDDFSIHAAHVVCRELGYVEAKSWTASSSYG 115

Qy     319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSD----GADLELRLADG 374
Db     116 KGEGPIWLDNLHCTGNEATLAACTSNGWGVTDCKHTEDVGVVCSDKRIPGFKFDNSLINQ 175

Qy     375 SNNCS----------------------GRVEVRIHEQWWTICDQNWKNEQALVVCKQLG 411
Db     176 IENLNIQVEDIRIRAILSTYRKRTPVMEGYVEVKEGKTWKQICDKHWTAKNSRVVCGMFG 235

Qy     412 CP---------FSVFGSRRAKPSNEARDIWINSISCTGNESALWDC------TYDGKAKR 456
Db     236 FPGERTYNTKVYKMFASRR------KQRYWPFSMDCTGTEAHISSCKLGPQVSLDPMKNV 289
```

```
Qy    457 TCFRRSDAGVIC---------------SDKADLDL-RLVGAHSPCYGRLEVKYQGEWGT 499
           ||     | | |               : | :  | || |       ||:||   |||||
Db    290 TCENGLPAVVSCVPGQVFSPDGPSRFRKAYKPEQPLVRLRGGAYIGEGRVEVLKNGEWGT 349

Qy    500 VCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSG 559
           || |:|   :|:|||::|| |           :    :  ||| |::: | ||| :| ||: :
Db    350 VCDDKWDLVSASVVCRELGFGSAKEAVTGSRLGQGIGPIHLNEIQCTGNEKSIIDCKFNA 409

Qy    560 WGKHNCVHREDVIVTCSGDATW---GLRLVGGSNRCSGRLEVYFQGR----WGTVCDDGW 612
               | | ||  | |: |          ||| || |    ||:||  :          || ||   |
Db    410 -ESQGCNHEEDAGVRCNTPAMGLQKKLRLNGGRNPYEGRVEVLVERNGSLVWGMVCGQNW 468

Qy    613 NSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIW-----------LDDVSCDGDESDLW 660
                | ||| ||        ||| ||   : : |                :  | | | |  |
Db    469 GIVEAMVVCRQL---------GLGFASNAFQETWYWHGDVNSNKVVMSGVKCSGTELSLA 519

Qy    661 SCRNSGWGNNDCSHSEDV----------GVICSD-ASDMEL--RLVGGSSRCAGKVEVN 706
            ||: |          |||          || ||: | |: |    :|  ::       :
Db    520 HCRHDG---------EDVACPQGGVQYGAGVACSETAPDLVLNAEMVQQTTYLEDRPMFM 570

Qy    707 VQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMS---NSGCT--- 760
           :|      ||                    :  | ||        :|     | |     |   |:| :
Db    571 LQ------CA-----------MEENCLSASAAQTDPTTGYRRLLRFSSQIHNNGQSDFR 612

Qy    761 ---GGEASLW-DCIR-WEWKQTACH---LNMEASLICSAHRQPRLVGADMPCSGRVEVKH 812
              |   | :| || | :    :   |    ||:  : :    |:     :   |  | | :: :
Db    613 PKNGRHAWIWHDCHRHYHSMEVFTHYDLLNLNGTKVAEGHKASFCL-EDTECEGDIQKNY 671

Qy    813 ADTWRSVC----DSDFSLHAANVLCRELNCG--DAISLSVGDH 849
                  |     |    ::    ::    :::|    |    :   ||:
Db    672 E------CANFGDQGITMGCWDMYRHDIDCQWVDITDVPPGDY 708
```

RESULT 12
US-09-448-076-7
; Sequence 7, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: MNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-448-076-7

```
  Query Match             7.5%;  Score 607.5;  DB 4;  Length 774;
  Best Local Similarity   27.9%;  Pred. No. 5.6e-45;
  Matches  196;  Conservative   86;  Mismatches  258;  Indels  163;  Gaps   28;

Qy    260 LELRLVGGTNR-CMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
           ::||| |    :     ||||:    |:|||||    ::   || |||::||    |   :
Db     56 IQLRLAGQKRKHSEGRVEVYYDGQWGTVCDDDFSIHAAHVVCRELGYVEAKSWTASSSYG 115

Qy    319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSD----GADLELRLADG 374
            |     :|||  : |:|||: |   |   :|       || |   || |:|||       |     :   | :
Db    116 KGEGPIWLDNLHCTGNEATLAACTSNGWGVTDCKHTEDVGVVCSDKRIPGFKFDNSLINQ 175

Qy    375 SNNCS----------------------GRVEVRIHEQWWTICDQNWKNEQALVVCKQLG 411
            | :                           | |||:   :  |   |||::|    :  : |||    |
Db    176 IENLNIQVEDIRIRAILSTYRKRTPVMEGYVEVKEGKTWKQICDKHWTAKNSRVVCGMFG 235

Qy    412 CP---------FSVFGSRRAKPSNEARDIWINSISCTGNESALWDC------TYDGKAKR 456
           |            :  :| |||          :   |   |:  |||  |:  :     |           :   |
Db    236 FPGERTYNTKVYKMFASRR------KQRYWPFSMDCTGTEAHISSCKLGPQVSLDPMKNV 289

Qy    457 TCFRRSDAGVIC---------------SDKADLDL-RLVGAHSPCYGRLEVKYQGEWGT 499
           ||     | | |               : | :  | || |       ||:||   |||||
Db    290 TCENGLPAVVSCVPGQVFSPDGPSRFRKAYKPEQPLVRLRGGAYIGEGRVEVLKNGEWGT 349

Qy    500 VCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSG 559
           || |:|   :|:|||::|| |           :    :  ||| |::: | ||| :| ||: :
Db    350 VCDDKWDLVSASVVCRELGFGSAKEAVTGSRLGQGIGPIHLNEIQCTGNEKSIIDCKFNA 409

Qy    560 WGKHNCVHREDVIVTCSGDATW---GLRLVGGSNRCSGRLEVYFQGR----WGTVCDDGW 612
               | | ||  | |: |          ||| || |    ||:||  :          || ||   |
Db    410 -ESQGCNHEEDAGVRCNTPAMGLQKKLRLNGGRNPYEGRVEVLVERNGSLVWGMVCGQNW 468

Qy    613 NSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIW-----------LDDVSCDGDESDLW 660
                | ||| ||        ||| ||   : : |                :  | | | |  |
Db    469 GIVEAMVVCRQL---------GLGFASNAFQETWYWHGDVNSNKVVMSGVKCSGTELSLA 519

Qy    661 SCRNSGWGNNDCSHSEDV----------GVICSD-ASDMEL--RLVGGSSRCAGKVEVN 706
            ||: |          |||          || ||: | |: |    :|  ::       :
Db    520 HCRHDG---------EDVACPQGGVQYGAGVACSETAPDLVLNAEMVQQTTYLEDRPMFM 570

Qy    707 VQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMS---NSGCT--- 760
           :|      ||                    :  | ||        :|     | |     |   |:| :
Db    571 LQ------CA-----------MEENCLSASAAQTDPTTGYRRLLRFSSQIHNNGQSDFR 612

Qy    761 ---GGEASLW-DCIR-WEWKQTACH---LNMEASLICSAHRQPRLVGADMPCSGRVEVKH 812
              |   | :| || | :    :   |    ||:  : :    |:     :   |  | | :: :
Db    613 PKNGRHAWIWHDCHRHYHSMEVFTHYDLLNLNGTKVAEGHKASFCL-EDTECEGDIQKNY 671

Qy    813 ADTWRSVC----DSDFSLHAANVLCRELNCG--DAISLSVGDH 849
                  |     |    ::    ::    :::|    |    :   ||:
Db    672 E------CANFGDQGITMGCWDMYRHDIDCQWVDITDVPPGDY 708
```

RESULT 13
US-09-702-572-7
; Sequence 7, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/702,572
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-572-7

```
  Query Match             7.5%;  Score 607.5;  DB 4;  Length 774;
  Best Local Similarity   27.9%;  Pred. No. 5.6e-45;
  Matches  196;  Conservative   86;  Mismatches  258;  Indels  163;  Gaps   28;

Qy    260 LELRLVGGTNR-CMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQ 318
           ::||| |    :     ||||:    |:|||||    ::   || |||::||    |   :
Db     56 IQLRLAGQKRKHSEGRVEVYYDGQWGTVCDDDFSIHAAHVVCRELGYVEAKSWTASSSYG 115

Qy    319 SGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSD----GADLELRLADG 374
            |     :|||  : |:|||: |   |   :|       || |   || |:|||       |     :   | :
Db    116 KGEGPIWLDNLHCTGNEATLAACTSNGWGVTDCKHTEDVGVVCSDKRIPGFKFDNSLINQ 175

Qy    375 SNNCS----------------------GRVEVRIHEQWWTICDQNWKNEQALVVCKQLG 411
            | :                           | |||:   :  |   |||::|    :  : |||    |
Db    176 IENLNIQVEDIRIRAILSTYRKRTPVMEGYVEVKEGKTWKQICDKHWTAKNSRVVCGMFG 235

Qy    412 CP---------FSVFGSRRAKPSNEARDIWINSISCTGNESALWDC------TYDGKAKR 456
```

```
Db     236 FPGERTYNTKVYKMFASRR------KQRYWPFSMDCTGTEAHISSCKLGPQVSLDPMKNV 289

Qy     457 TCFRRSDAGVIC---------------SDKADLDL-RLVGAHSPCYGRLEVKYQGEWGT 499

Db     290 TCENGLPAVVSCVPGQVFSPDGPSRFRKAYKPEQPLVRLRGGAYIGEGRVEVLKNGEWGT 349

Qy     500 VCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSG 559

Db     350 VCDDKWDLVSASVVCRELGFGSAKEAVTGSRLGQGIGPIHLNEIQCTGNEKSIIDCKFNA 409

Qy     560 WGKHNCVHREDVIVTCSGDATW---GLRLVGGSNRCSGRLEVYFQGR----WGTVCDDGW 612

Db     410 -ESQGCNHEEDAGVRCNTPAMGLQKKLRLNGGRNPYEGRVEVLVERNGSLVWGMVCGQNW 468

Qy     613 NSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIW-----------LDDVSCDGDESDLW 660

Db     469 GIVEAMVVCRQL---------GLGFASNAFQETWYWHGDVNSNKVVMSGVKCSGTELSLA 519

Qy     661 SCRNSGWGNNDCSHSEDV-----------GVICSD-ASDMEL--RLVGGSSRCAGKVEVN 706

Db     520 HCRHDG---------EDVACPQGGVQYGAGVACSETAPDLVLNAEMVQQTTYLEDRPMFM 570

Qy     707 VQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMS---NSGCT--- 760

Db     571 LQ------CA------------MEENCLSASAAQTDPTTGYRRLLRFSSQIHNNGQSDFR 612

Qy     761 ---GGEASLW-DCIR-WEWKQTACH---LNMEASLICSAHRQPRLVGADMPCSGRVEVKH 812

Db     613 PKNGRHAWIWHDCHRHYHSMEVFTHYDLLNLNGTKVAEGHKASFCL-EDTECEGDIQKNY 671

Qy     813 ADTWRSVC----DSDFSLHAANVLCRELNCG--DAISLSVGDH 849

Db     672 E------CANFGDQGITMGCWDMYRHDIDCQWVDITDVPPGDY 708

RESULT 14
US-09-034-916-15
; Sequence 15, Application US/09034916
; Patent No. 6046314
;  GENERAL INFORMATION:
;    APPLICANT:  GEBE, JOHN A.
;    APPLICANT:  SIADAK, ANTHONY W.
;    APPLICANT:  ARUFFO, ALEJANDRO A.
;    TITLE OF INVENTION:  SpALPHA: A NOVEL SCAVENGER RECEPTOR
;    TITLE OF INVENTION:  CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
;    TITLE OF INVENTION:  ANTIBODIES THERETO
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBINS & ASSOCIATES
;      STREET:  90 MIDDLEFIELD ROAD, SUITE 200
;      CITY:  MENLO PARK
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,916
;      FILING DATE:  04-MAR-1998
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/039,956
;      FILING DATE:  06-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  5998-0020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 325-7812
;      TELEFAX:  (650) 325-7823
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  100 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-034-916-15

  Query Match             5.7%;  Score 462;  DB 3;  Length 100;
  Best Local Similarity   77.0%;  Pred. No. 1.9e-33;
  Matches   77;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     1037 RLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSG 1096

Db        1 RLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTG 60

Qy     1097 PIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICS 1136

Db       61 PIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICS 100

RESULT 15
US-09-034-916-17
; Sequence 17, Application US/09034916
; Patent No. 6046314
;  GENERAL INFORMATION:
;    APPLICANT:  GEBE, JOHN A.
;    APPLICANT:  SIADAK, ANTHONY W.
;    APPLICANT:  ARUFFO, ALEJANDRO A.
;    TITLE OF INVENTION:  SpALPHA: A NOVEL SCAVENGER RECEPTOR
;    TITLE OF INVENTION:  CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
;    TITLE OF INVENTION:  ANTIBODIES THERETO
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBINS & ASSOCIATES
;      STREET:  90 MIDDLEFIELD ROAD, SUITE 200
;      CITY:  MENLO PARK
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,916
;      FILING DATE:  04-MAR-1998
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/039,956
;      FILING DATE:  06-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  5998-0020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 325-7812
;      TELEFAX:  (650) 325-7823
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  100 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-034-916-17

  Query Match             5.3%;  Score 432;  DB 3;  Length 100;
  Best Local Similarity   71.0%;  Pred. No. 8.7e-31;
  Matches   71;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;
```

```
Qy   1247 RVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTG 1306
          |:: | | ||||||||| ||||||||||||| :|:||||||||| || | ::| ||||||
Db      1 RLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTG 60

Qy   1307 TIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCS 1346
           |||::::|||||| |||| |: || |:||||||| | |:
Db     61 PIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCT 100
```

Search completed: May 12, 2003, 06:44:15
Job time : 31 secs

```
OM protein - protein search, using sw model

Run on:          May 12, 2003, 06:37:26 ; Search time 151 Seconds
                                           (without alignments)
                                           885.519 Million cell updates/sec

Title:           US-09-759-130B-381
Perfect score:   8138
Sequence:        1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters:     349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                 1:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                 2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                 3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                 4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                 5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                 6:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                 7:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                 8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 9:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                 10:  /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
                 11:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                 12:  /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
                 13:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                 14:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8138	100.0	1453	9	US-10-042-431-11	Sequence 11, Appl
2	8138	100.0	1453	9	US-09-759-130B-381	Sequence 381, App
3	7913	97.2	1413	9	US-10-042-431-13	Sequence 13, Appl
4	7913	97.2	1413	9	US-09-759-130B-383	Sequence 383, App
5	7426	91.3	1319	9	US-10-042-431-14	Sequence 14, Appl
6	7426	91.3	1319	9	US-09-759-130B-384	Sequence 384, App
7	3441.5	42.3	1151	9	US-09-977-577-13	Sequence 13, Appl
8	3429	42.1	1116	9	US-09-977-577-10	Sequence 10, Appl
9	3429	42.1	1156	9	US-09-977-577-12	Sequence 12, Appl
10	3396.5	41.7	1149	9	US-09-977-577-11	Sequence 11, Appl
11	3082	37.9	1436	9	US-10-042-431-78	Sequence 78, Appl
12	3082	37.9	1436	9	US-09-759-130B-448	Sequence 448, App
13	822.5	10.1	822	9	US-09-147-947-6	Sequence 6, Appli
14	719.5	8.8	761	9	US-09-147-947-4	Sequence 4, Appli
15	691.5	8.5	347	9	US-09-905-291A-148	Sequence 148, App
16	691.5	8.5	347	9	US-09-902-853-148	Sequence 148, App
17	691.5	8.5	347	9	US-09-907-824-148	Sequence 148, App
18	691.5	8.5	347	9	US-09-907-841-148	Sequence 148, App
19	691.5	8.5	347	9	US-09-904-011-148	Sequence 148, App
20	691.5	8.5	347	9	US-09-906-742-148	Sequence 148, App
21	691.5	8.5	347	9	US-09-906-838-148	Sequence 148, App
22	691.5	8.5	347	9	US-09-907-613-148	Sequence 148, App
23	691.5	8.5	347	9	US-09-907-942-148	Sequence 148, App
24	691.5	8.5	347	9	US-10-227-884-114	Sequence 114, App
25	691.5	8.5	347	9	US-09-904-820-148	Sequence 148, App
26	691.5	8.5	347	9	US-09-904-859-148	Sequence 148, App
27	691.5	8.5	347	9	US-09-909-204-148	Sequence 148, App
28	691.5	8.5	347	9	US-10-230-163-114	Sequence 114, App
29	691.5	8.5	347	9	US-09-904-786-148	Sequence 148, App
30	691.5	8.5	347	9	US-09-906-646-148	Sequence 148, App
31	691.5	8.5	347	9	US-09-906-700-148	Sequence 148, App
32	691.5	8.5	347	9	US-09-902-903-148	Sequence 148, App
33	691.5	8.5	347	9	US-09-903-749A-148	Sequence 148, App
34	691.5	8.5	347	9	US-09-903-786-148	Sequence 148, App
35	691.5	8.5	347	9	US-10-218-631-114	Sequence 114, App
36	691.5	8.5	347	9	US-10-230-338-114	Sequence 114, App
37	691.5	8.5	347	9	US-09-902-736-148	Sequence 148, App
38	691.5	8.5	347	9	US-09-904-119-148	Sequence 148, App
39	691.5	8.5	347	9	US-09-904-956-148	Sequence 148, App
40	691.5	8.5	347	9	US-09-907-794-148	Sequence 148, App
41	691.5	8.5	347	9	US-10-230-414-114	Sequence 114, App
42	691.5	8.5	347	9	US-09-902-692-148	Sequence 148, App
43	691.5	8.5	347	9	US-09-903-520-148	Sequence 148, App
44	691.5	8.5	347	9	US-09-903-943-148	Sequence 148, App
45	691.5	8.5	347	9	US-09-904-462-148	Sequence 148, App

ALIGNMENTS

```
RESULT 1
US-10-042-431-11
; Sequence 11, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:  SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 1453
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-431-11

  Query Match             100.0%;  Score 8138;  DB 9;  Length 1453;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1453;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60

Qy         61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120

Qy        121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180
```

```
Qy    181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240

Qy    241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300

Qy    301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360

Qy    361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420

Qy    421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480

Qy    481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540

Qy    541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600

Qy    601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660

Qy    661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720

Qy    721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780

Qy    781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840

Qy    841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900

Qy    901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960

Qy    961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020

Qy   1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080

Qy   1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140

Qy   1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200

Qy   1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260

Qy   1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320

Qy   1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380

Qy   1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440

Qy   1441 SLLGVLPASEATK 1453
          |||||||||||||
Db   1441 SLLGVLPASEATK 1453


RESULT 2
US-09-759-130B-381
; Sequence 381, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 381
;   LENGTH: 1453
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-130B-381

  Query Match             100.0%;  Score 8138;  DB 9;  Length 1453;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1453;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1 MMLPQNSWHIDFGRCCCHQNLFSAVVTCILLLNSCFLISSFNGTDLELRLVNGDGPCSGT 60

Qy      61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120

Qy     121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTIC 180

Qy     181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWG 240

Qy     241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 NHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVC 300

Qy     301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 KQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVI 360

Qy     361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSR 420

Qy     421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 RAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVG 480

Qy     481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWL 540

Qy     541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 DDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYF 600

Qy     601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 QGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLW 660

Qy     661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 SCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWG 720

Qy     721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 MNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACH 780

Qy     781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 LNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGD 840

Qy     841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDV 900

Qy     901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSV 960

Qy     961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 RVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLS 1020

Qy    1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCG 1080

Qy    1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 VAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTA 1140

Qy    1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 LRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKT 1200

Qy    1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVE 1260

Qy    1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320

Qy    1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 FLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHLALILSSIFGLLLLVLFILFL 1380

Qy    1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGCEDASDT 1440

Qy    1441 SLLGVLPASEATK 1453
           |||||||||||||
Db    1441 SLLGVLPASEATK 1453


RESULT 3
US-10-042-431-13
; Sequence 13, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:  SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 1413
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-431-13

  Query Match             97.2%;  Score 7913;  DB 9;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      41 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 60

Qy     101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 120

Qy     161 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 180

Qy     221 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 240

Qy     281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 300

Qy       341 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 360

Qy       401 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 420

Qy       461 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 480

Qy       521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 540

Qy       581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 600

Qy       641 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 660

Qy       701 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 720

Qy       761 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 780

Qy       821 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 840

Qy       881 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 900

Qy       941 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960

Qy      1001 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1020

Qy      1061 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1080

Qy      1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1140

Qy      1181 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1200

Qy      1241 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1260

Qy      1301 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHL 1360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHL 1320

Qy      1361 ALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDP 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDP 1380

Qy      1421 HGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1453
             |||||||||||||||||||||||||||||||||
Db      1381 HGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1413
```

RESULT 4

```
US-09-759-130B-383
; Sequence 383, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383
;   LENGTH: 1413
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-130B-383

  Query Match             97.2%;  Score 7913;  DB 9;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        41 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 60

Qy       101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 120

Qy       161 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 180
```

```
Qy    221 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 240

Qy    281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 300

Qy    341 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 360

Qy    401 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 420

Qy    461 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 480

Qy    521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 540

Qy    581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 600

Qy    641 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 660

Qy    701 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 720

Qy    761 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 780

Qy    821 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 840

Qy    881 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 900

Qy    941 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960

Qy   1001 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1020

Qy   1061 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1080

Qy   1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1140

Qy   1181 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1200

Qy   1241 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1260

Qy   1301 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHL 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGHL 1320

Qy   1361 ALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDP 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ALILSSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDP 1380

Qy   1421 HGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1453
          |||||||||||||||||||||||||||||||||
Db   1381 HGTRTSDDTPNHGCEDASDTSLLGVLPASEATK 1413
```

```
RESULT 5
US-10-042-431-14
; Sequence 14, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:  SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 1319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-431-14

  Query Match             91.3%;  Score 7426;  DB 9;  Length 1319;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1319;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     41 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 60

Qy    101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 120

Qy    161 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 180

Qy    221 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 240

Qy    281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 300

Qy    341 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 360

Qy    401 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 420

Qy    461 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     421 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 480

Qy     521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 540

Qy     581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 600

Qy     641 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 660

Qy     701 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 720

Qy     761 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 780

Qy     821 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 840

Qy     881 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 900

Qy     941 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960

Qy    1001 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1020

Qy    1061 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1080

Qy    1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1140

Qy    1181 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1200

Qy    1241 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1260

Qy    1301 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1319
```

RESULT 6
US-09-759-130B-384
; Sequence 384, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 384
;   LENGTH: 1319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-130B-384

```
  Query Match             91.3%;  Score 7426;  DB 9;  Length 1319;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1319;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      41 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCKQLGCPFSFAMFRF 60

Qy     101 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GQAVTRHGKIWLDDVSCYGNESALWECQHREWGSHNCYHGEDVGVNCYGEANLGLRLVDG 120

Qy     161 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSSFISSGVVNSPAVLRPIWLD 180

Qy     221 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQ 240

Qy     281 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWD 300

Qy     341 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKN 360

Qy     401 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 EQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFR 420

Qy     461 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 RSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCG 480
```

```
Qy    521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDAT 540

Qy    581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNAST 600

Qy    641 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCA 660

Qy    701 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCT 720

Qy    761 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVC 780

Qy    821 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 DSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHP 840

Qy    881 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 EDTCIHSREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQ 900

Qy    941 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960

Qy   1001 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTI 1020

Qy   1061 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGW 1080

Qy   1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIV 1140

Qy   1181 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWI 1200

Qy   1241 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDAS 1260

Qy   1301 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 FGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSLNASSGH 1319
```

RESULT 7
US-09-977-577-13
; Sequence 13, Application US/09977577
; Patent No. US20020155995A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USES T
; FILE REFERENCE: MOESTRUP-1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 1151
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-977-577-13

  Query Match 42.3%; Score 3441.5; DB 9; Length 1151;
  Best Local Similarity 55.2%; Pred. No. 9e-239;
  Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

```
Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db    581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db    641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db    701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db    761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db    821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940
```

```
Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDDTPNHGC 1434
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADDLDLMNSSEN--SHES 1115

Qy   1435 EDASDTSLLGV---LPASEATK 1453
Db   1116 ADFSAAELISVSKFLPISGMEK 1137
```

RESULT 8
US-09-977-577-10
; Sequence 10, Application US/09977577
; Patent No. US20020155995A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USES T
; FILE REFERENCE: MOESTRUP=1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 1116
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-977-577-10

```
  Query Match             42.1%;  Score 3429;  DB 9;  Length 1116;
  Best Local Similarity   56.2%;  Pred. No. 6.8e-238;
  Matches  598;  Conservative  161;  Mismatches  294;  Indels   12;  Gaps    4;

Qy    364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
Db     41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy    424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
Db    101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy    483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db    161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy    543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
Db    221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy    603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
Db    281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy    663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
Db    341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy    723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
Db    401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy    783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
Db    461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy    843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
Db    521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy    903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
Db    581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy    963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db    641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy   1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db    701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy   1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
Db    761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy   1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
Db    821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy   1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db    881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy   1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
Db    941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy   1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
Db   1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy   1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
Db   1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102
```

RESULT 9
US-09-977-577-12
; Sequence 12, Application US/09977577
; Patent No. US20020155995A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USE
; FILE REFERENCE: MOESTRUP=1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 1156
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-977-577-12

```
  Query Match             42.1%;  Score 3429;  DB 9;  Length 1156;
  Best Local Similarity   56.2%;  Pred. No. 7.1e-238;
```

```
  Matches  598;  Conservative  161;  Mismatches  294;  Indels  12;  Gaps  4;

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
           | | |||| || | |||||||:: |:| |:|:   |   |     |:| |||||  ::      |
Db      41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy     424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
             |  :   ||::  :|| ||||||||| :||   |  :  |   :  |||| ||| ::|::||
Db     101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy     483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
           : |  ||:|:|:|| |||||  | ::   :|:|:|:|:|| ||    :      | :  |  | ||||| ||
Db     161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy     543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
           : | |||| :|:|:| ||||||| | ||  | ||  |      |||| |    ||||||| |||
Db     221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy     603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
            |||:|||||:|   ||| | || ||:::   :|   ||| |:| |||| ||| | |   :| |
Db     281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy     663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
           ::  ||  :  |:|:|| || ||| ||:|||| || ||||| ||| :|   :| :|   |||:
Db     341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy     723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
            |:|||||| ||||::  |  :   :   :   :   |  |  |  |  ||||| |:|      |
Db     401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy     783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
            || :  |||||:||||| |:||||||||| ||| |:||||||| ||:|||||| ||    :
Db     461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy     843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRL 902
            |:   | |||:||| |||:||||| |:||:|||: . || || |||:||||||||||::|||
Db     521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSRYTEIRL 580

Qy     903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRV 962
           ||||: |:|:||:  || |||||::|||  |||  ||||:|| || |||| ||   |: : ::
Db     581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKGNGQI 640

Qy     963 WGHRFHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
            | | ||| |  |    : :| :|:|||  |       ||||: :|  :|  | ::     |
Db     641 WRHMFHCTGTEQHMGDCPVTALGASLCPSEQVASVICSGNQSQTLSSCNSSSLGPTRPTI 700

Qy    1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
            || ||: |:|   :||||:|  |||||||||||:| ||||||| |||||||||||||:||||| |
Db     701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEA 760

Qy    1083 FNATVSAHFGEGSGPIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142
            ||| ||||||:||||||:: | | || :|||  |  ||||  :||||||||||||| :||
Db     761 INATGSAHFGEGTGPIWLDEMKCNGKESRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLR 820

Qy    1143 LYSETETESCAGRLEVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGS 1202
            | ||    |:||||||||||| ||:||: :::          |:|||||| : | :: | | |  |
Db     821 LTSEASREACAGRLEVFYNGAWGTVGKSSMSETTVGVVCRQLGCADKGKINPASLDKAMS 880

Qy    1203 GFMWVDDIQCPKTHISIWQCLSAPWERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
            ||||::||||    ::||| |:|||:|::|:||||||:::||:: | | ||||||||
Db     881 IPMWVDNVQCPKGPDTLWQCPSSPWEKRLASPSEETWITCDNKIRLQEGPTSCSGRVEIW 940

Qy    1263 HAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNESFL 1322
            | |||||||||||| :|:|||||||| ||  | :::| |||||| |||::::|||||| |
Db     941 HGGSWGTVCDDSWDLDDAQVVCQQLGCGPALKAFKEAEFGQGTGPIWLNEVKCKGNESSL 1000

Qy    1323 WDCHAKPWGQSDCGHKEDAGVRCSGQSLK------SLNASSGHLALILSSIFGLLLLVLF 1376
            ||| |: || |:||||||| | |:  |::       :  ||  : |  |.|::|| :|
Db    1001 WDCPARRWGHSECGHKEDAAVNCTDISVQKTPQKATTGRSSRQSSFIAVGILGVVLLAIF 1060

Qy    1377 I--LFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKRED 1419
            :   |||  | |:|:   | ||:|    :  :  || :||   :|
Db    1061 VALFFLTKKRRQRQR---LAVSSRGENLVHQIQYREMNSCLNADD 1102
```

```
RESULT 10
US-09-977-577-11
; Sequence 11, Application US/09977577
; Patent No. US20020155995A1
; GENERAL INFORMATION:
;  APPLICANT: MOESTRUP, Soren
;  APPLICANT:  MOLLER, Holger J.
;  TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USE
;  FILE REFERENCE: MOESTRUP=1A
;  CURRENT APPLICATION NUMBER: US/09/977,577
;  CURRENT FILING DATE:  2001-10-16
;  PRIOR APPLICATION NUMBER: US 60/270,120
;  PRIOR FILING DATE: 2001-02-22
;  PRIOR APPLICATION NUMBER: DK PA 2001 00039
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: DK PA 2000 01543
;  PRIOR FILING DATE: 2000-10-16
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 1149
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-977-577-11

  Query Match             41.7%;  Score 3396.5;  DB 9;  Length 1149;
  Best Local Similarity   54.4%;  Pred. No. 1.5e-235;
  Matches  597;  Conservative  161;  Mismatches  295;  Indels  45;  Gaps  5;

Qy     364 GADLELRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAK 423
           | | |||| || | |||||||:: |:| |:|:   |   |     |:| |||||  ::      |
Db      41 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVICNQLGCPTAIKAPGWAN 100

Qy     424 PSNEARDIWINSISCTGNESALWDCTYDGKAKRT-CFRRSDAGVICSDKADLDLRLVGAH 482
             |  :   ||::  :|| ||||||||| :||   |  :  |   :  |||| ||| ::|::||
Db     101 SSAGSGRIWMDHVSCRGNESALWDCKHDGWGKHSNCTHQQDAGVTCSDGSNLEMRLTRGG 160

Qy     483 SPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
           : |  ||:|:|:|| |||||  | ::   :|:|:|:|:|| ||    :      | :  |  | ||||| ||
Db     161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFSGSSNFGEGSGPIWFDD 220

Qy     543 VSCIGNESNIWDCEHSGWGKHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQG 602
           : | |||| :|:|:| ||||||| | ||  | ||  |      |||| |    ||||||| |||
Db     221 LICNGNESALWNCKHQGWGKHNCDHAEDAGVICSKGADLSLRLVDGVTECSGRLEVRFQG 280

Qy     603 RWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTGYGKIWLDDVSCDGDESDLWSC 662
            |||:|||||:|   ||| | || ||:::   :|   ||| |:| |||| ||| | |   :| |
Db     281 EWGTICDDGWDSYDAAVACKQLGCPTAVTAIGRVNASKGFGHIWLDSVSCQGHEPAVWQC 340

Qy     663 RNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMN 722
           ::  ||  :  |:|:|| || ||| ||:|||| || ||||| ||| :|   :| :|   |||:
Db     341 KHHEWGKHYCNHNEDAGVTCSDGSDLELRLRGGGSRCAGTVEVEIQRLLGKVCDRGWGLK 400

Qy     723 IAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLN 782
            |:|||||| ||||::  |  :   :   :   :   |  |  |  |  ||||| |:|      |
Db     401 EADVVCRQLGCGSALKTSYQVYSKIQATNTWLFLSSCNGNETSLWDCKNWQWGGLTCDHY 460

Qy     783 MEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAI 842
            || :  |||||:||||| |:||||||||| ||| |:||||||| ||:|||||| ||    :
Db     461 EEAKITCSAHREPRLVGGDIPCSGRVEVKHGDTWGSICDSDFSLEAASVLCRELQCGTVV 520

Qy     843 SLSVGDHFGKGNGLTWAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSR------ 896
            |:   | |||:||| |||:||||| |:||:|||:   || || |||:||||||
Db     521 SILGGAHFGEGNGQIWAEEFQCEGHESHLSLCPVAPRPEGTCSHSRDVGVVCSSKTQKTS 580

Qy     897 --------------------------YTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHW 929
                                     ||::|||||: |:|:||:  || |||||::||
Db     581 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNSHW 640

Qy     930 DPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLLDNCQMTVLGAPPC 989
            | ||| |||||:|| || ||||  ||  |: : ::| | ||| |  :  : :| :|:|||  |
```

```
Db    641 DIEDAHVLCQQLKCGVALSTPGGARFGKGNGQIWRHMFHCTGTEQHMGDCPVTALGASLC 700

Qy    990 IHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRV 1049
Db    701 PSEQVASVICSGNQSQTLSSCNSSSLGPTRPTIPEESAVACIESGQLRLVNGGGRCAGRV 760

Qy   1050 EIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSGPIWLDDLNCTGTE 1109
Db    761 EIYHEGSWGTICDDSWDLSDAHVVCRQLGCGEAINATGSAHFGEGTGPIWLDEMKCNGKE 820

Qy   1110 SHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSVGR 1169
Db    821 SRIWQCHSHGWGQQNCRHKEDAGVICSEFMSLRLTSEASREACAGRLEVFYNGAWGTVGK 880

Qy   1170 RNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTHISIWQCLSAPWER 1229
Db    881 SSMSETTVGVVCRQLGCADKGKINPASLDKAMSIPMWVDNVQCPKGPDTLWQCPSSPWEK 940

Qy   1230 RISSPAEETWITCEDRIRVRGGDTECSGRVEIWHAGSWGTVCDDSWDLAEAEVVCQQLGC 1289
Db    941 RLASPSEETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDDSWDLDDAQVVCQQLGC 1000

Qy   1290 GSALAALRDASFGQGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQS 1349
Db   1001 GPALKAFKEAEFGQGTGPIWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCTDIS 1060

Qy   1350 LK------SLNASSGHLALILSSIFGLLLLVLFI--LFLTWCRVQKQKHLPLRVSTRRRG 1401
Db   1061 VQKTPQKATTGRSSRQSSFIAVGILGVVLLAIFVALFFLTKKRRQRQR---LAVSSRGEN 1117

Qy   1402 SLEENLFHEMETCLKRED 1419
Db   1118 LVHQIQYREMNSCLNADD 1135
```

RESULT 11
US-10-042-431-78
; Sequence 78, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 1436
; TYPE: PRT
; ORGANISM: Bos sp.
US-10-042-431-78

```
  Query Match            37.9%;  Score 3082;  DB 9;  Length 1436;
  Best Local Similarity  41.7%;  Pred. No. 7.9e-213;
  Matches  613;  Conservative  208;  Mismatches  495;  Indels  154;  Gaps  23;

Qy     28 CILLLNSCFLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCK 87
Db     13 CVLLLG-----TMVGGQALELRLKDGVHRCEGRVEVKHQGEWGTVDGYRWTLKDASVVCR 67

Qy     88 QLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQH---REWGSHNCYHGE 141
Db     68 QLGCGAAIGFPGGAYFGPGL---GPIWLLYTSCEGTESTVSDCEHSNIKDYRNDGYNHGR 124

Qy    142 DVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS 201
Db    125 DAGVVCSG----FVRLAGGDGPCSGRVEVHSGEAWIPVSDGNFTLATAQIICAELGCGKA 180

Qy    202 FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE 261
Db    181 VSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPCPGGTCHHSGSAQVVCSAYSEVR 240

Qy    262 LRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAADVVCKQLGCGTALHFAGLPHLQSGS 321
Db    241 L-MTNGSSQCEGQVEMNISGQWRALCASHWSLANANVICRQLGCGVAISTPGGPHLVEEG 299

Qy    322 DVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGR 381
Db    300 DQILTARFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGN----------------- 342

Qy    382 VEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGN 441
Db    343 -QIQVLPQ---------------------CNDSV-----SQPTGSAAS----------- 363

Qy    442 ESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVC 501
Db    364 -------------------EDSAPYCSDSR--QLRLVDGGGPCAGRVEILDQGSWGTIC 401

Qy    502 HDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWG 561
Db    402 DDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKESHVWRCPSRGWG 461

Qy    562 KHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVC 621
Db    462 QHNCRHKQDAGVICS--EFLALRMVSEDQQCAGWLEVFYNGTWGSVCRNPMEDITVSTIC 519

Qy    622 SQLDCPSSIIGMGLGNAST----GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
Db    520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWNYNSCSPKEE 575

Qy    678 VGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAI 737
Db    576 AYIWCADS--RQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVCKQLGCGEAL 633

Qy    738 RVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRL 797
Db    634 DATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAGVICSGF--VRL 691

Qy    798 VGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLT 857
Db    692 AGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESDGQV 751

Qy    858 WAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLV-NGKSQCDGQVEIN 916
Db    752 WAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMKNGTSQCEGQVEMK 811

Qy    917 VLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLL 976
Db    812 ISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQISTAQFHCSGAESFL 871

Qy    977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036
Db    872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSAASEESSPYCSDSRQL 931

Qy   1037 RLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSG 1096
Db    932 RLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQLGCGEALNATGSAHFGAGSG 991

Qy   1097 PIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRL 1156
Db    992 PIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVICSEFLALRMVS--EDQQCAGWL 1049

Qy   1157 EVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTH 1216
Db   1050 EVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTSVGLREGSRPRWVDLIQCRKMD 1109

Qy   1217 ISIWQCLSAPWERRISSPAEETWITCE---------------DRIRVRGGDTECSGRVE 1260
```

```
                 |:||| | ||:     || || :|:||                 :::|:||||:|||||||
Db     1110 TSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTAACTDREKLRLRGGDSECSGRVE 1169

Qy     1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
             :|| |||||||||| ||||||||||||| || |:| |:|| | |:||||:::| |'||
Db     1170 VWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEAVRSAAFGPGNGSIWLDEVQCGGRES 1229

Qy     1321 FLWDCHAKPWGQSDCGHKEDAGVRCSG----------------QSLKSLNASSGHLALIL 1364
             |||| |:||||||| |:|||||||||                ||  : : .| | |||
Db     1230 SLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLPTTTAGTRTTSNSLPGIFSLPGVLCLIL 1289

Qy     1365 SSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTR 1424
             |:  |:|::|      | | | ::       |   :    :| | :: |::  | :::  |:
Db     1290 GSLLFLVLVILVTQLLRW-RAER------RALSSYEDALAEAVYEELDYLLTQKEGLGSP 1342

Qy     1425 TS-DDTPNHGCEDASDTSLLGVLPASEATK 1453
                 | |:   :|| :  : |    |:  :
Db     1343 DQMTDVPDENYDDAEEVPVPGTPSPSQGNE 1372
```

RESULT 12
US-09-759-130B-448
; Sequence 448, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 448
;   LENGTH: 1436
;   TYPE: PRT
;   ORGANISM: Bovine
US-09-759-130B-448

```
  Query Match             37.9%;  Score 3082;  DB 9;  Length 1436;
  Best Local Similarity   41.7%;  Pred. No. 7.9e-213;
  Matches  613;  Conservative  208;  Mismatches  495;  Indels  154;  Gaps   23;

Qy       28 CILLLNSCFLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDGWNTTASTVVCK 87
            |:|||      :   |  |||||  :|   | |  ||||  ||:||||     |     ::|||:
Db       13 CVLLLG-----TMVGGQALELRLKDGVHRCEGRVEVKHQGEWGTVDGYRWTLKDASVVCR 67

Qy       88 QLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQH---REWGSHNCYHGE 141
            ||||  :          ||  :    | |||    || | ||  : :|:|    ::: :    ||
Db       68 QLGCGAAIGFPGGAYFGPGL---GPIWLLYTSCEGTESTVSDCEHSNIKDYRNDGYNHGR 124

Qy      142 DVGVNCYGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLNTAAVVCRQLGCPSS 201
            | || | |        :||   |:   ||||||   | |   :   |   :  | || ::| :|||   :
Db      125 DAGVVCSG----FVRLAGGDGPCSGRVEVHSGEAWIPVSDGNFTLATAQIICAELGCGKA 180

Qy      202 FISSGVVNSPAVLRPIWLDDILCQGNELALWNCRHRGWGNHDCSHNEDVTLTCYDSSDLE 261
                 |                :| ::  |:| |  || |          | |:     :  |   |::
Db      181 VSVLGHELFRESSAQVWAEEFRCEGEEPELWVCPRVPCPGGTCHHSGSAQVVCSAYSEVR 240

Qy      262 LRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCGTALHFAGLPHLQSGS 321
            | :   |:::| |:||:  | |:|   :|    |:  | |:|:|:||||| |:     | |||
Db      241 L-MTNGSSQCEGQVEMNISGQWRALCASHWSLANANVICRQLGCGVAISTPGGPHLVEEG 299

Qy      322 DVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICSDGADLELRLADGSNNCSGR 381
            | :         ||| ||||| |    :           || |  |    |||||
Db      300 DQILTARFHCSGAESFLWSCPVTALGGPDCSHGNTASVICSGN---------------- 342

Qy      382 VEVRIHEQWWTICDQNWKNEQALVVCKQLGCPFSVFGSRRAKPSNEARDIWINSISCTGN 441
             ::::   |                              |  ||          ::|:  |
Db      343 -QIQVLPQ---------------------CNDSV-----SQPTGSAAS---------- 363

Qy      442 ESALWDCTYDGKAKRTCFRRSDAGVICSDKADLDLRLVGAHSPCYGRLEVKYQGEWGTVC 501
                                     |:    |||       ||||        || ||:|:   || |||:|
Db      364 -------------------EDSAPYCSDSR--QLRLVDGGGPCAGRVEILDQGSWGTIC 401

Qy      502 EDRWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWG 561
              | |     :| |||:|||||: ::   |    :|     |||||||:::| | ||::| |    |||
Db      402 DDGWDLDDARVVCRQLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKESHVWRCPSRGWG 461

Qy      562 KHNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVC 621
            :||| |::|  | ||       ·||:|     :|:| |||:: | ||:|| :              : :|
Db      462 QHNCRHKQDAGVICS--EFLALRMVSEDQQCAGWLEVFYNGTWGSVCRNPMEDITVSTIC 519

Qy      622 SQLDCPSSIIGMGLGNAST----GYGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
             || |        | |:|       |:   |:| :  || | :    |  ||  |     :
Db      520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSLWQCPSDPWNYNSCSPKEE 575

Qy      678 VGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAI 737
                : |:|:      ::|||  |  ||:|:||:  ||: |: :| : | :: | ||||:|| || |:
Db      576 AYIWCADS--RQIRLVDGGGRCSGRVEILDQGSWGTICDDRWDLDDARVVCKQLGCGEAL 633

Qy      738 EVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRL 797
                :        |     :    | :       | | |: :| |       | | |:|     |:       :| :|||          ||
Db      634 DATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSWGWRQHNCNHQEDAGVICSGF--VRL 691

Qy      798 VGADMPCSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGNGLT 857
              | | ||||||||       :  |     | | | :|:|    | |:| || || |:|:          | : :|
Db      692 AGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESDGQV 751

Qy      858 WAEKFQCEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTDVRLV-NGKSQCDGQVEIN 916
            |||:|:|:|  |     |     ||  |   |    |:||    ||||  ||:|:|:|  || ||||:|||:
Db      752 WAEEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCSVYTEVQLMKNGTSQCEGQVEMK 811

Qy      917 VLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGNESLL 976
            : | | :|| :||      :| | |:|||| || || |  | :  :|    ::      :||| | || |
Db      812 ISGRWRALCASHWSLANANVVCRQLGCGVAISTPRGPHLVEGGDQISTAQFHCSGAESFL 871

Qy      977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036
             :| :| || | |  ||||  ||||:|: || |    |     :| |   ||   | |:   | : ·:|
Db      872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSAASEESSPYCSDSRQL 931

Qy     1037 RLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCQKLGCGVAFNATVSAHFGEGSG 1096
            |||||    | |||||    | ||||||| ||| || |||::|||| | ||| ||||| |||
Db      932 RLVDGGGPCGGRVEILDQGSWGTICDDDWDLDDARVVCRQLGCGEALNATGSAHFGAGSG 991
```

```
Qy   1097 PIWLDDLNCTGTESHLWQCPSRGWGQHDCRHKEDAGVICSEFTALRLYSETETESCAGRL 1156
Db    992 PIWLDDLNCTGKESHVWRCPSRGWGRHDCRHKEDAGVICSEFLALRMVS--EDQQCAGWL 1049

Qy   1157 EVFYNGTWGSVGRRNITTAIAGIVCRQLGCGENGVVSLAPLSKTGSGFMWVDDIQCPKTH 1216
Db   1050 EVFYNGTWGSVCRSPMEDITVSVICRQLGCGDSGSLNTSVGLREGSRPRWVDLIQCRKMD 1109

Qy   1217 ISIWQCLSAPWERRISSPAEETWITCE---------------DRIRVRGGDTECSGRVE 1260
Db   1110 TSLWQCPSGPWKYSSCSPKEEAYISCEGRRPKSCPTAAACTDREKLRLRGGDSECSGRVE 1169

Qy   1261 IWHAGSWGTVCDDSWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKGNES 1320
Db   1170 VWHNGSWGTVCDDSWSLAEAEVVCQQLGCGQALEAVRSAAFGPGNGSIWLDEVQCGGRES 1229

Qy   1321 FLWDCHAKPWGQSDCGHKEDAGVRCSG---------------QSLKSLNASSGHLALIL 1364
Db   1230 SLWDCVAEPWGQSDCKHEEDAGVRCSGVRTTLPTTTAGTRTTSNSLPGIFSLPGVLCLIL 1289

Qy   1365 SSIFGLLLLVLFILFLTWCRVQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTR 1424
Db   1290 GSLLFLVLVILVTQLLRW-RAER------RALSSYEDALAEAVYEELDYLLTQKEGLGSP 1342

Qy   1425 TS-DDTPNHGCEDASDTSLLGVLPASEATK 1453
Db   1343 DQMTDVPDENYDDAEEVPVPGTPSPSQGNE 1372
```

RESULT 13
US-09-147-947-6
; Sequence 6, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoko
; APPLICANT: YAMAGUCHI, No. US20020160490A1omi
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-6

```
  Query Match             10.1%;  Score 822.5;  DB 9;  Length 822;
  Best Local Similarity   29.6%;  Pred. No. 7.7e-51;
  Matches  229;  Conservative   96;  Mismatches  307;  Indels  141;  Gaps   26;

Qy    239 WGNHDCSHNEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADV 298
Db    107 WGYCDCRHG-------------SVRLRGGKNEFEGTVEVYASGVWGTVCSSHWDDSDASV 153

Qy    299 VCKQL-----GCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLH 353
Db    154 ICHQLQLGGKGIAKQTPFSGL-----GLIPIYWSNVRCRGDEENILLCEKDIWQGGVCPQ 208

Qy    354 QNDVSVIC--SDGADLE-LRLADGSNNCSGRVEVRIHEQWWTICDQNWKNEQALVVCKQL 410
Db    209 KMAAAVTCSFSHGPTFPIIRLAGGSSVHEGRVELYHAGQWGTVCDDQWDDADAEVICRQL 268

Qy    411 GCPFSVFGSRRAKPSNEARDIWINSISCTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470
Db    269 GLSGIAKAWHQAYFGEGSGPVMLDEVRCTGNELSIEQCPKSSWGEHNCGHKEDAGVSCTP 328

Qy    471 KADLDLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRNAAVVCKQLG--CGKPMHVFGM 528
Db    329 LTDGVIRLAGGKGSHEGRLEVYYRGQWGTVCDDGWTELNTYVVCRQLGFKYGKQA---SA 385

Qy    529 TYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKHNCVHREDVIVTC--SGDA-----TW 581
Db    386 NHFEESTGPIWLDDVSCSGKETRFLQCSRRQWGRHDCSHREDVSIACYPGGEGHRLSLGF 445

Qy    582 GLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQLDCPSSIIGMGLGNASTG 641
Db    446 PVRLMDGENKKEGRVEVFINGQWGTICDDGWTDKDAAVICRQLGYKGPARARTMAYFGEG 505

Qy    642 YGKIWLDDVSCDGDESDLWSCRNSGWGNNDCSHSEDVGVICSDASDMELRLVGGSSRCAG 701
Db    506 KGPIHVDNVKCTGNERSLADCIKQDIGRHNCRHSEDAGVIC----DYFGKKASGNS---- 557

Qy    702 KVEVNVQGAVGILCANGWGMNIAEVVCRQLECGSAIRVSREPHFTERTLHILMSNSGCTG 761
Db    558 ----NKESLSSV------------------CGLRL-------LHRRQKRIIGGKNSLRG 587

Qy    762 GEASLWDCIRWEWKQTACHLNMEASLICSAHRQPRLVGADMPCSGRVEVKHADTWRSVCD 821
Db    588 G---------WPW-QVSLRLK-------SSHGDGRLL-----CGA-----------TLLS 614

Qy    822 SDFSLHAANVLCRELNCGDAISLSVGDH-------FGKGNG---LTWAEKFQCEGSETHL 871
Db    615 SCWVLTAAHCFKRYGNSTRSYAVRVGDYHTLVPEEFEEEIGVQQIVIHREYRPDRSDYDI 674

Qy    872 ALCPIVQHPEDTCIH-SREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWGSLCDTHWD 930
Db    675 ALVRL-QGPEEQCARFSSHVLPACLPLWRERPQKTASNC-------YITGWG-------- 718

Qy    931 PEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN---ESLLDNCQ 980
Db    719 -DTGRAYSRTLQ-QAAIPLLPKRFCEERYKGRFTGRMLCAGNLHEHKRVDSCQ 769
```

RESULT 14
US-09-147-947-4
; Sequence 4, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoko
; APPLICANT: YAMAGUCHI, No. US20020160490A1omi
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 761
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-4

```
  Query Match              8.8%;  Score 719.5;  DB 9;  Length 761;
  Best Local Similarity   31.8%;  Pred. No. 1.8e-43;
  Matches  191;  Conservative   80;  Mismatches  233;  Indels   97;  Gaps   19;

Qy    451 DGKAKRTCFRRSDAGVI------CSDKADLD-LRLVGAHSPCYGRLEVKYQGEWGTVCHD 503
Db    134 DGSGRPWCFYRNAQGKVDWGYCDCGQGPALPVIRLVGGNSGHEGRVELYHAGQWGTICDD 193
```

```
Qy    504 RWSTRNAAVVCKQLGCGKPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCEHSGWGKH 563
           :|   :| |:|:|||       :     :| |  |||| ||:| | ||| :|  |   | ||:|
Db    194 QWDNADADVICRQLGLSGIAKAWHQAHFGEGSGPILLDEVRCTGNELSIEQCPKSSWGEH 253

Qy    564 NCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNSKAAAVVCSQ 623
           || |:|| |:|        :|| || :    |||||::|:||||||||       | |
Db    254 NCGHKEDAGVSCVPLTDGVIRLAGGKSTHEGRLEVYYKGQWGTVCDDGWTEMNTYVAC-- 311

Qy    624 LDCPSSIIGMGLGNAST-----GYGK-IWLDDVSCDGDESDLWSCRNSGWGNNDCSHSED 677
               ::|   |  |:     |  : |||||||| | |     |     || :|||| ||
Db    312 -----RLLGFKYGKQSSVNHFDGSNRPIWLDDVSCSGKEVSFIQCSRRQWGRHDCSHRED 366

Qy    678 VGVICSDASD-------MELRLVGGSSRCAGKVEVNVQGAVGILCANGWGMNIAEVVCRQ 730
           ||: |   ||         :||| | ::   |:||| | |   | :| :||    | |:||
Db    367 VGLTCYPDSDGHRLSPGFPIRLVDGENKKEGRVEVFVNGQWGTICDDGWTDKHAAVICRQ 426

Qy    731 LECGSAIRVSREPHFTERTLHILMSNSGCTGGEASLWDCIRWEWKQTACHLNMEASLICS 790
           |      |      :| |     | | |   ||| | :| ||:: :  :  |   : :| :||
Db    427 LGYKGPARARTMAYFGEGKGPIHMDNVKCTGNEKALADCVKQDIGRHNCRHSEDAGVICD 486

Qy    791 -----------------------AHRQPRLVGADMPCSG--------RVEVKHADTWRS 818
                                  || |::| :    |          |:   | | |
Db    487 YLEKKASSSGNKEMLSSGCGLRLLHRRQKRIIGGNNSLRGAWPWQASLRLRSAHGD-GRL 545

Qy    819 VC-----DSDFSLHAANVLCRELNCGDAISLSVGDH-------FGKGNG---LTWAEKFQ 863
           :|        | : | ||:     |   |    : :: |||:          | :   |     :    ::
Db    546 LCGATLLSSCWVLTAAHCFKRYGNNSRSYAVRVGDYHTLVPEEFEQEIGVQQIVIHRNYR 605

Qy    864 CEGSETHLALCPIVQHPEDTCIH-SREVGVVCSRYTDVRLVNGKSQCDGQVEINVLGHWG 922
            : |:  :||   : | | : |    |    |      |       | |         :: | ||
Db    606 PDRSDYDIALVRL-QGPGEQCARLSTHVLPACLPLWRERPQKTASNC------HITG-WG 657

Qy    923 SLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVRVWGHRFHCLGN---ESLLDNC 979
                :  |   | |   |:    ::  ||   ::  |   | | ||   :: :|:|
Db    658 ---------DTGRAYSRTLQ-QAAVPLLPKRFCKERYKGLFTGRMLCAGNLQEDNRVDSC 707

Qy    980 Q 980
          |
Db    708 Q 708
```

RESULT 15
US-09-905-291A-148
; Sequence 148, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 148
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-148

```
  Query Match             8.5%;  Score 691.5;  DB 9;  Length 347;
  Best Local Similarity   41.3%;  Pred. No. 6.7e-42;
  Matches  148;  Conservative   40;  Mismatches  139;  Indels   31;  Gaps    9;

Qy     21 LFSAVVTCILLLNSC----FLISSFNGTDLELRLVNGDGPCSGTVEVKFQGQWGTVCDDG 76
           |||      |:|  |      || |        :||| |    | | |||: :|||||||||
Db      4 LFS------LILAICTRPGFLASPSG-----VRLVGGLHRCEGRVEVEQKGQWGTVCDDG 52

Qy     77 WNTTASTVVCKQLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYGNESALWECQHREWG 133
           |:       |:|::|||   :          :     :  |: :   ||| |  |    | :|:  |
Db     53 WDIKDVAVLCRELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEE-- 110

Qy    134 SHNCYHGEDVGVNC------YGEANLGLRLVDGNNSCSGRVEVKFQERWGTICDDGWNLN 187
            ::| | || | :|          :        |:|| ||    | |||||| | :| |:|   ||:|
Db    111 VYDCSHDEDAGASCENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLR 170

Qy    188 TAAVVCRQLGCPSSFISSGVVNSPAVLR-PIWLDDILCQGNELALWNCRHRGWGNHDCSH 246
            | ||||||||   : ::       |     |   | ||||   :  |  |  |   | :|     || : |:|
Db    171 AAKVVCRQLGCGRAVLTQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNH 230

Qy    247 NEDVTLTCYDSSDLELRLVGGTNRCMGRVELKIQGRWGTVCHHKWNNAAADVVCKQLGCG 306
           :||   :  |  |   |    ||||||  |  |  ||:|:   :| ||:||        |        ||||||||
Db    231 DEDTWVECEDPFD--LRLVGGDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCG 288

Qy    307 TAL--HFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSGTVNFDCLHQNDVSVICS 362
            :|    |        |    :||| | ||| |   |   |:|           || ||  ||:||||
Db    289 KSLSPSFRDRKCYGPGVGRIWLDNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICS 346
```

Search completed: May 12, 2003, 06:47:08
Job time : 162 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         May 12, 2003, 10:35:25 ; Search time 254 Seconds
                                          (without alignments)
                                          7108.859 Million cell updates/sec

Title:          US-09-759-130B-381
Perfect score:  8138
Sequence:       1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters:      1567708

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09759130/runat_06052003_123555_23406/app_query.fasta_1.1607
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09759130_@CGN_1_1_158_@runat_06052003_123555_23406
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8138	100.0	4359	9	US-10-042-431-10	Sequence 10, Appl
2	8138	100.0	4359	9	US-09-759-130B-380	Sequence 380, App
3	8138	100.0	4628	9	US-10-042-431-9	Sequence 9, Appli
4	8138	100.0	4628	9	US-09-759-130B-379	Sequence 379, App
5	3082	37.9	4308	9	US-10-042-431-79	Sequence 79, Appl
6	3082	37.9	4308	9	US-09-759-130B-449	Sequence 449, App
7	1142.5	14.0	4344	10	US-09-917-800A-1712	Sequence 1712, Ap
8	830	10.2	2562	9	US-09-147-947-5	Sequence 5, Appli
9	818	10.1	473	9	US-09-918-995-32438	Sequence 32438, A
10	722	8.9	2614	9	US-09-147-947-3	Sequence 3, Appli
11	691.5	8.5	1686	9	US-09-905-291A-147	Sequence 147, App
12	691.5	8.5	1686	9	US-09-902-853-147	Sequence 147, App
13	691.5	8.5	1686	9	US-09-907-824-147	Sequence 147, App
14	691.5	8.5	1686	9	US-09-907-841-147	Sequence 147, App
15	691.5	8.5	1686	9	US-09-904-011-147	Sequence 147, App
16	691.5	8.5	1686	9	US-09-906-742-147	Sequence 147, App
17	691.5	8.5	1686	9	US-09-906-838-147	Sequence 147, App
18	691.5	8.5	1686	9	US-09-907-613-147	Sequence 147, App
19	691.5	8.5	1686	9	US-09-907-942-147	Sequence 147, App
20	691.5	8.5	1686	9	US-10-227-884-113	Sequence 113, App
21	691.5	8.5	1686	9	US-09-904-820-147	Sequence 147, App
22	691.5	8.5	1686	9	US-09-904-859-147	Sequence 147, App
23	691.5	8.5	1686	9	US-09-909-204-147	Sequence 147, App
24	691.5	8.5	1686	9	US-10-230-163-113	Sequence 113, App
25	691.5	8.5	1686	9	US-09-904-786-147	Sequence 147, App
26	691.5	8.5	1686	9	US-09-906-646-147	Sequence 147, App
27	691.5	8.5	1686	9	US-09-906-700-147	Sequence 147, App
28	691.5	8.5	1686	9	US-09-902-903-147	Sequence 147, App
29	691.5	8.5	1686	9	US-09-903-749A-147	Sequence 147, App
30	691.5	8.5	1686	9	US-09-903-786-147	Sequence 147, App
31	691.5	8.5	1686	9	US-10-218-631-113	Sequence 113, App
32	691.5	8.5	1686	9	US-10-230-338-113	Sequence 113, App
33	691.5	8.5	1686	9	US-09-902-736-147	Sequence 147, App
34	691.5	8.5	1686	9	US-09-904-119-147	Sequence 147, App
35	691.5	8.5	1686	9	US-09-904-956-147	Sequence 147, App
36	691.5	8.5	1686	9	US-09-907-794-147	Sequence 147, App
37	691.5	8.5	1686	9	US-10-230-414-113	Sequence 113, App
38	691.5	8.5	1686	9	US-09-902-692-147	Sequence 147, App
39	691.5	8.5	1686	9	US-09-903-520-147	Sequence 147, App
40	691.5	8.5	1686	9	US-09-903-943-147	Sequence 147, App
41	691.5	8.5	1686	9	US-09-904-462-147	Sequence 147, App
42	691.5	8.5	1686	9	US-09-905-056-147	Sequence 147, App
43	691.5	8.5	1686	9	US-09-907-925-147	Sequence 147, App
44	691.5	8.5	1686	9	US-09-904-553-147	Sequence 147, App
45	691.5	8.5	1686	9	US-09-905-381-147	Sequence 147, App

ALIGNMENTS

```
RESULT 1
US-10-042-431-10
; Sequence 10, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:  SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4359
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-042-431-10

Alignment Scores:
Pred. No.:             0           Length:        4359
```

```
Score:                    8138.00      Matches:        1453
Percent Similarity:       100.00%      Conservative:   0
Best Local Similarity:    100.00%      Mismatches:     0
Query Match:              100.00%      Indels:         0
DB:                       9            Gaps:           0

US-09-759-130B-381 (1-1453) x US-10-042-431-10 (1-4359)

Qy        1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 60

Qy       21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 120

Qy       41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180

Qy       61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240

Qy       81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 300

Qy      101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 360

Qy      121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 420

Qy      141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 480

Qy      161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 540

Qy      181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600

Qy      201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 660

Qy      221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 720

Qy      241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 780

Qy      261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 840

Qy      281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GGAAGGTGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 900

Qy      301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
            ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db      901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 960

Qy      321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1020

Qy      341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1080

Qy      361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1140

Qy      381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1200

Qy      401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1260

Qy      421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1320

Qy      441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1380

Qy      461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1440

Qy      481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1500

Qy      501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1560

Qy      521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620

Qy      541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1680

Qy      561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1740

Qy      581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1800

Qy      601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860

Qy      621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1920

Qy      641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 1980

Qy      661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2040

Qy      681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2041 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100

Qy    701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2160

Qy    721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220

Qy    741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2280

Qy    761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2340

Qy    781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2400

Qy    801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2460

Qy    821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2520

Qy    841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2580

Qy    861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2640

Qy    881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2700

Qy    901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2760

Qy    921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2820

Qy    941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2880

Qy    961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940

Qy    981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000

Qy   1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3060

Qy   1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3120

Qy   1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3180

Qy   1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3240

Qy   1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3300

Qy   1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3360

Qy   1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3420

Qy   1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480

Qy   1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3540

Qy   1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3600

Qy   1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3660

Qy   1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3720

Qy   1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3780

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3840

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3900

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3960

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4020

Qy   1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4080

Qy   1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4140

Qy   1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4200

Qy   1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4260
```

```
Qy     1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4261 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320

Qy     1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
            |||||||||||||||||||||||||||||||||||||||
Db     4321 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4359

RESULT 2
US-09-759-130B-380
; Sequence 380, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 380
;   LENGTH: 4359
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-130B-380

Alignment Scores:
Pred. No.:              0              Length:         4359
Score:                  8138.00        Matches:        1453
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     9              Gaps:           0

US-09-759-130B-381 (1-1453) x US-09-759-130B-380 (1-4359)

Qy        1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 60

Qy       21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 120

Qy       41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180

Qy       61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240

Qy       81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 300

Qy      101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 360

Qy      121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 420

Qy      141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 480

Qy      161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 540

Qy      181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600

Qy      201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 660

Qy      221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 720

Qy      241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 780

Qy      261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 840

Qy      281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 900

Qy      301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 960

Qy      321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1020

Qy      341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1080

Qy      361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1140

Qy      381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1200

Qy      401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1260

Qy      421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1320

Qy      441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1380

Qy      461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1440

Qy      481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1500

Qy      501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1560

Qy      521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620

Qy      541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1680

Qy      561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1740

Qy      581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1800

Qy      601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860

Qy      621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1920

Qy      641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 1980

Qy      661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2040

Qy      681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100

Qy      701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2160

Qy      721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220

Qy      741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2280

Qy      761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2340

Qy      781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2400

Qy      801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2460

Qy      821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2520

Qy      841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2580

Qy      861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2640

Qy      881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2700

Qy      901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2760

Qy      921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2820

Qy      941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2880

Qy      961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940

Qy      981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000

Qy     1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3060

Qy     1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3120

Qy     1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3180

Qy     1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3240

Qy     1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3300

Qy     1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3360
```

```
Qy   1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3420

Qy   1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480

Qy   1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3540

Qy   1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3600

Qy   1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3660

Qy   1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3720

Qy   1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3780

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3840

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3900

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3960

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4020

Qy   1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4080

Qy   1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4140

Qy   1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4200

Qy   1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4260

Qy   1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320

Qy   1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
          |||||||||||||||||||||||||||||||||||||||
Db   4321 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4359
```

RESULT 3
US-10-042-431-9
; Sequence 9, Application US/10042431
; Publication No. US20020182675A1

```
; GENERAL INFORMATION:
;  APPLICANT: MCCARTHY, Sean A
;  APPLICANT:  BARNES, Thomas M
;  APPLICANT:  FRASER, Christopher C
;  APPLICANT:   SHARP, John D
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;  TITLE OF INVENTION:  PREVENTIVE, THERAPEUTIC, AND OTHER USES
;  FILE REFERENCE: 10147-6U2
;  CURRENT APPLICATION NUMBER: US/10/042,431
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 4628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-042-431-9

Alignment Scores:
Pred. No.:              0               Length:         4628
Score:                  8138.00         Matches:        1453
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-759-130B-381 (1-1453) x US-10-042-431-9 (1-4628)

Qy      1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     28 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 87

Qy     21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 147

Qy     41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207

Qy     61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 267

Qy     81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 327

Qy    101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 387

Qy    121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 447

Qy    141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 507

Qy    161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 567

Qy    181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 627

Qy    201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 687

Qy     221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 747

Qy     241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 807

Qy     261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     808 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 867

Qy     281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     868 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 927

Qy     301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 987

Qy     321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     988 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1047

Qy     341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1048 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1107

Qy     361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1108 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1167

Qy     381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1168 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1227

Qy     401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1228 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1287

Qy     421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1288 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1347

Qy     441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1407

Qy     461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1467

Qy     481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1468 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1527

Qy     501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1528 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1587

Qy     521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1588 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1647

Qy     541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1648 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1707

Qy     561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1708 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767

Qy     581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1768 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1827

Qy     601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1828 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887

Qy     621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1947

Qy     641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1948 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2007

Qy     661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2008 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067

Qy     681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2068 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127

Qy     701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2128 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2187

Qy     721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2188 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2247

Qy     741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2248 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2307

Qy     761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2308 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2367

Qy     781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2368 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2427

Qy     801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2428 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2487

Qy     821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2547

Qy     841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2548 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2607

Qy     861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2608 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2667

Qy     881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727

Qy     901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2728 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787

Qy     921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2847
```

```
Qy      941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2907

Qy      961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967

Qy      981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2968 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027

Qy     1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3087

Qy     1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3088 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3147

Qy     1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3207

Qy     1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267

Qy     1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3327

Qy     1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3328 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3387

Qy     1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3388 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3447

Qy     1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3448 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507

Qy     1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3508 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3567

Qy     1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3568 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3627

Qy     1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3628 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3687

Qy     1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3688 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3747

Qy     1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3748 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3807

Qy     1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3808 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3867

Qy     1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3927

Qy     1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3928 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3987

Qy     1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3988 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4047

Qy     1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4048 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4107

Qy     1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4108 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4167

Qy     1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4168 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4227

Qy     1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4228 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4287

Qy     1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4288 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4347

Qy     1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
            |||||||||||||||||||||||||||||||||||||||
Db     4348 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4386
```

```
RESULT 4
US-09-759-130B-379
; Sequence 379, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION:  NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:   PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:   USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
```

```
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 379
;   LENGTH: 4628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-130B-379

Alignment Scores:
Pred. No.:               0            Length:        4628
Score:                   8138.00      Matches:       1453
Percent Similarity:      100.00%      Conservative:  0
Best Local Similarity:   100.00%      Mismatches:    0
Query Match:             100.00%      Indels:        0
DB:                      9            Gaps:          0

US-09-759-130B-381 (1-1453) x US-09-759-130B-379 (1-4628)

Qy      1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     28 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 87

Qy     21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 147

Qy     41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207

Qy     61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 267

Qy     81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 327

Qy    101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 387

Qy    121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 447

Qy    141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 507

Qy    161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 567

Qy    181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 627

Qy    201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 687

Qy    221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 747

Qy    241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 807

Qy    261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 867

Qy    281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 927

Qy    301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 987

Qy    321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    988 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1047

Qy    341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1048 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1107

Qy    361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1167

Qy    381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1227

Qy    401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1287

Qy    421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1347

Qy    441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1407

Qy    461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1467

Qy    481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1527

Qy    501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1587

Qy    521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1588 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1647

Qy    541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1648 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1707

Qy    561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1708 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767

Qy    581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1768 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1827

Qy    601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1828 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887

Qy    621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1947

Qy    641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1948 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2007

Qy    661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2008 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067

Qy    681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2068 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127

Qy    701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2128 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2187

Qy    721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2188 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2247

Qy    741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2248 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2307

Qy    761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2308 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2367

Qy    781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2368 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2427

Qy    801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2428 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2487

Qy    821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2547

Qy    841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2548 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2607

Qy    861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2608 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2667

Qy    881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727

Qy    901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2728 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787

Qy    921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2847

Qy    941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2907

Qy    961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967

Qy    981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2968 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027

Qy   1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3087

Qy   1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3088 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3147

Qy   1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3207

Qy   1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267

Qy   1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3327

Qy   1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3328 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3387

Qy   1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3447

Qy   1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507

Qy   1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3567

Qy   1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3627

Qy   1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3628 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3687

Qy   1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3688 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3747

Qy   1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3748 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3807

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3808 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3867

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3927

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3928 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3987

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3988 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4047

Qy   1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4048 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4107
```

```
Qy   1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4108 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4167

Qy   1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4168 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4227

Qy   1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4228 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4287

Qy   1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4288 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4347

Qy   1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
          |||||||||||||||||||||||||||||||||||||||
Db   4348 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4386
```

RESULT 5
US-10-042-431-79
; Sequence 79, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 4308
;   TYPE: DNA
;   ORGANISM: Bos sp.
US-10-042-431-79

```
Alignment Scores:
Pred. No.:              2.74e-302      Length:        4308
Score:                  3082.00        Matches:       613
Percent Similarity:     55.85%         Conservative:  208
Best Local Similarity:  41.70%         Mismatches:    495
Query Match:            37.87%         Indels:        154
DB:                     9              Gaps:          23

US-09-759-130B-381 (1-1453) x US-10-042-431-79 (1-4308)

Qy     28 CysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu 47
          |||:::|||||||||                       :::      |||      |||||
Db     37 TGTGTCCTCCTCCTCGGC---------------ACCATGGTGGGTGGTCAAGCTCTGGAG 81

Qy     48 LeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
          ||||||||    :::|||          |||   |||   ||||||||||||   ||||||
Db     82 CTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCACCAAGGA 141

Qy     68 GlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrValValCysLys 87
          :::||||||||||||||              |||                ::::::|||||||||:::
Db    142 GAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGTGTGCAGA 201

Qy     88 GlnLeuGlyCysProPheSerPheAlaMet---------PheArgPheGlyGlnAlaVal 104
          ||||||||||||        :::                     ||||||         :::
Db    202 CAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTTGGGCCAGGACTT 261

Qy    105 ThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeu 124
                   |||   |||||||||            ||||||   |||   ||||||   :::
Db    262 ---------GGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTC 312

Qy    125 TrpGluCysGlnHis---------ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141
          :::|||:::|||             :::::::::   :::            ||||||
Db    313 AGTGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGG 372

Qy    142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161
          |||   ||||||   |||   |||                   :::||||||   |||:::
Db    373 GATGCTGGAGTAGTCTGCTCAGGA------------TTTGTGCGTCTGGCTGGAGGGGAT 420

Qy    162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
                |||||||||||||||||||||            |||   |||      :::   |||
Db    421 GGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGAT 480

Qy    182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
                :::   |||   ||||||   ::::::|||   :::|||||||||         :::
Db    481 GGAACTTCACACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCT 540

Qy    202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
                      |||                                 :::|||   ::::::
Db    541 GTGTCTGTCCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAG 600

Qy    222 IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241
                |||:::|||   |||         ||||||   |||
Db    601 TTCAGGTGTGAGGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGG 660

Qy    242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu 261
                |||   |||:::            :::   |||            |||::::::
Db    661 GGCACGTGTCACCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGAAGTCCGG 720

Qy    262 LeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281
          |||   :::         |||:::::::::|||   |||::::::||||||:::   |||   |||
Db    721 CTC---ATGACAAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGA 777

Qy    282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLys 301
          :::|||         :::|||               |||:::   |||   |||:::|||:::|||:::
Db    778 CAATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGT 837

Qy    302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321
          |||||||||||||||   |||:::            |||   |||||||||
Db    838 CAGCTCGGCTGTGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGT 897

Qy    322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
          |||   :::                     |||||||||   |||||||||||||||   |||
Db    898 GATCAGATCCTAACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGT 957

Qy    342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
                :::                  ||||||   |||   |||      ||||||||||||
Db    958 CCTGTGACTGCCCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGC 1017

Qy    362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
          |||
Db   1018 TCAGGAAAC--------------------------------------------------- 1026

Qy    382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
             ::::::::::::      |||
Db   1027 ---CAGATCCAGGTGCTTCCCCAG------------------------------------ 1047

Qy    402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
                                       |||      ||||||
Db   1048 ----------------------------TGCAACGACTCCGTG----------------- 1062

Qy    422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
          ::::::|||:::        |||
Db   1063 TCTCAACCTACAGGCTCTGCGGCCTCA--------------------------------- 1089

Qy    442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461

Db   1089 ------------------------------------------------------------ 1089

Qy    462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
             |||:::         |||||||||                  ||||||||||||
```

```
Db   1090 GAGGACAGCGCCCCCTACTGCTCAGACAGCAGG------CAGCTCCGCCTGGTGGACGGG 1143

Qy    482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
                   ||||||   ||||||:::|||:::          ||||||   ||||||||||:::|||
Db   1144 GGCGGTCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGT 1203

Qy    502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
             |||   |||          :::|||   ||||||||||:::||||||||||||||||:::
Db   1204 GATGACGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAA 1263

Qy    522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
             ::::::         |||      :::|||      ||||||||||||||||||||||||
Db   1264 GCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGAC 1323

Qy    542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
          :::::::::|||   |||   ||||||:::::::|||   |||          |||||||||
Db   1324 AACTTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGG 1383

Qy    562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
          :::|||||||||   |||::::::|||        |||   ||||||
Db   1384 CAGCACAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCA------GAGTTCCTG 1437

Qy    582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601
             ||||||:::|||         :::|||:::|||   ||||||||||||::::::
Db   1438 GCCCTCAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAAT 1497

Qy    602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCys 621
          |||   ||||||:::||||||                     :::   :::|||
Db   1498 GGGACCTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGC 1557

Qy    622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
             ||||||   |||         |||            |||         |||:::|||
Db   1558 AGACAGCTTGGCTGTGGGGACAGT-----------GGAACCCTCAACTCTTCTGTTGCT 1605

Qy    641 ---------GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
                   |||:::         |||:::|||   :::   |||         ::::::
Db   1606 CTTAGAGAAGGTTTTAGGCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACC 1665

Qy    658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
             ||||||   |||   :::         |||         |||   ||||||      |||:::
Db   1666 TCTCTCTGGCAGTGTCCTTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAA 1725

Qy    678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697
             :::   |||:::|||:::         :::::::|||||||||   |||
Db   1726 GCCTATATCTGGTGTGCAGACAGC------AGACAGATCCGCCTGGTGGATGGAGGTGGT 1779

Qy    698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
          ||||||:::|||:::|||||||:::         ||||||:::   |||   :::|||   :::
Db   1780 CGCTGCTCTGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGAC 1839

Qy    718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIle 737
             |||   ::::::   |||   |||||||||::::|||||||  .||||||   |||:::
Db   1840 CGCTGGGACCTGGACGATGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTG 1899

Qy    738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
             :::               |||         :::      |||   :::
Db   1900 GACGCCACTGTCTCTTCCTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTG 1959

Qy    758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
             |||   |||   |||:::|||   :::|||   |||      |||:::|||
Db   1960 AACTGCAGAGGAGAGGAGTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACAC 2019

Qy    778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
             |||:::         :::|||   :::|||||||||                  ||||||
Db   2020 AACTGCAATCATCAAGAAGATGCAGGAGTCATCTGCTCAGGATTT------GTGCGTCTG 2073

Qy    798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
             |||   |||   ||||||||||||||||||||||||            :::   |||
Db   2074 GCTGGAGGAGATGGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACC 2133

Qy    818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
             |||   |||   :::|||:::|||         |||   |||:::|||   ||||||
Db   2134 CCAGTGTCTGATGGAAACTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGA 2193

Qy    838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
          ||||||   |||:::|||:::                  |||   :::   :::|||
Db   2194 TGTGGCAAGGCTGTGTCTGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTC 2253

Qy    858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
          ||||||||:::|||:::|||:::|||   |||         |||      ||||||   |||
Db   2254 TGGGCTGAAGAGTTCAGGTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTG 2313

Qy    878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
             |||      ||||||:::||||||               ||||||||||||   |||
Db   2314 CCCTGTCCAGGAGGCACATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTAC 2373

Qy    898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
          |||:::|||:::         ||||||   |||||||||:::||||||||||||:::
Db   2374 ACAGAAGTCCAGCTTATGAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAG 2433

Qy    917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
          :::   |||   |||   :::||||||   :::||||||         :::|||   |||
Db   2434 ATCTCTGGACGATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTT 2493

Qy    937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
          :::||||||||||||   ||||||   |||:::|||||||         |||   ::::::
Db   2494 GTCTGTCGTCAGCTCGGCTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTG 2553

Qy    957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
          |||         ::::::         :::|||||||||   |||   ||||||   |||
Db   2554 GAAGGAGGTGATCAGATCTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTG 2613

Qy    977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
             :::|||   :::|||      ||||||   |||   |||   ||||||||||||   |||
Db   2614 TGGAGTTGTCCTGTGACTGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCT 2673

Qy    997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
          ||||||||:::|||:::      ||||||   |||         |||         :::|||
Db   2674 GTGATCTGCTCAGGAAACCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAA 2733

Qy   1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
          |||      ||||||      |||   |||:::         |||   :::   ::::::|||
Db   2734 CCTGCAGGCTCTGCGGCCTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTC 2793

Qy   1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
          ||||||||||||||||||         |||   ||||||||||||||||||      |||
Db   2794 CGCCTGGTGGACGGGGGCGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCC 2853

Qy   1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
          ||||||||||||||||||||||   |||||||||   ||||||   ||||||||||::::::
Db   2854 TGGGGCACCATCTGTGATGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAG 2913

Qy   1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
          ||||||||||||   |||   |||||||||   |||||||||||||||   |||||||||
Db   2914 CTGGGCTGTGGAGAAGCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGG 2973

Qy   1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
          ||||||||||||||||||||||||||||||   |||||||||:::|||:::||||||
Db   2974 CCCATCTGGCTGGACGACCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCT 3033

Qy   1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
          ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db   3034 TCCCGGGGCTGGGGGCGGCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCA 3093

Qy   1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
          ||||||   ||||||||||:::   |||         :::   |||||||||      |||
Db   3094 GAGTTCCTGGCCCTCAGGATGGTGAGC------GAGGACCAGCAGTGTGCTGGGTGGCTG 3147

Qy   1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
          ||||||||||||||||||||||||||||||   |||         :::
Db   3148 GAGGTTTTCTACAACGGGACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACT 3207

Qy   1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
                :::::::||||||||||||||||||||||:::::::|||      ::::::   :::
Db   3208 GTGTCCGTGATCTGCAGACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTT 3267
```

```
Qy   1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
                :::   ||||||            |||||||||   |||||||||   |||
Db   3268 GGTCTCAGGGAAGGTTCTAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGAT 3327

Qy   1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
             |||:::||||||||   |||   ||||||:::                ||||||   |||
Db   3328 ACCTCTCTCTGGCAGTGTCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAG 3387

Qy   1237 GluThrTrpIleThrCysGlu------------------------------------- 1243
          |||   :::|||:::||||||
Db   3388 GAAGCCTACATCTCATGTGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGC 3447

Qy   1244 ---------AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
                   ::::::::|||:::||||||||||||:::|||||||||||||||||||
Db   3448 ACAGACAGAGAGAAGCTCCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAG 3507

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          :::||||||   |||||||||||||||||||||||||||||||   ||||||||||||
Db   3508 GTGTGGCACAACGGCTCCTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGCAGAGGCT 3567

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          |||||||||||||||||||||||||||||||   ||||||   |||:::|||   |||:::
Db   3568 GAGGTGGTGTGTCAGCAGCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCA 3627

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||   |||   |||:::|||||||||||||:::::::::|||   |||   ||||||
Db   3628 TTTGGCCCTGGAAATGGGAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCC 3687

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
            .||||||||||||   |||:::|||||||||||||||||||||   |||:::||||||
Db   3688 TCCCTGTGGGACTGTGTTGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGAT 3747

Qy   1341 AlaGlyValArgCysSerGly------------------------------------- 1347
          |||||||||||||||||||||
Db   3748 GCTGGTGTGAGGTGCTCTGGTGTAAGGACAACATTGCCCACGACCACAGCAGGGACCAGA 3807

Qy   1348 ---------GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeu 1364
                   ||||||      :::   :::      |||   |||   |||||||||
Db   3808 ACAACCTCAAATTCTCTCCCTGGCATCTTCTCCCTGCCTGGGGTTCTCTGCCTTATCCTG 3867

Qy   1365 SerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg 1384
             |||:::   |||:::|||:::::::|||            |||   |||   |||
Db   3868 GGGTCGCTTCTCTTCCTGGTCCTCGTCATCCTGGTGACTCAGCTACTCAGATGG---AGA 3924

Qy   1385 ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu 1404
             ::::::                  |||      :::         :::|||
Db   3925 GCAGAGCGC------------------AGAGCCTTATCCAGCTATGAAGATGCTCTTGCT 3966

Qy   1405 GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg 1424
          |||   ::::::   |||::::::      |||   :::::::::      |||:::
Db   3967 GAAGCTGTGTATGAGGAGCTCGATTACCTTCTGACACAGAAGGAAGGTCTGGGCAGCCCA 4026

Qy   1425 ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu 1443
                   |||   |||:::            :::||||||   :::      :::
Db   4027 GATCAGATGACTGATGTCCCTGATGAAAATTATGATGATGCTGAAGAAGTACCAGTGCCT 4086

Qy   1444 GlyValLeuProAlaSerGluAlaThrLys 1453
          |||            |||:::      :::
Db   4087 GGAACTCCTTCTCCCTCTCAGGGGAATGAG 4116
```

RESULT 6
US-09-759-130B-449
; Sequence 449, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Bovine
US-09-759-130B-449

```
Alignment Scores:
Pred. No.:                2.74e-302     Length:          4308
Score:                    3082.00       Matches:         613
Percent Similarity:       55.85%        Conservative:    208
Best Local Similarity:    41.70%        Mismatches:      495
Query Match:              37.87%        Indels:          154
DB:                       9             Gaps:            23
```

US-09-759-130B-381 (1-1453) x US-09-759-130B-449 (1-4308)

```
Qy     28 CysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu 47
          |||:::|||||||||                      :::         |||      |||||
Db     37 TGTGTCCTCCTCCTCGGC---------------ACCATGGTGGGTGGTCAAGCTCTGGAG 81

Qy     48 LeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
          ||||||||   :::|||         |||   |||   |||||||||||||   ||||||
Db     82 CTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCACCAAGGA 141

Qy     68 GlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrValValCysLys 87
          :::||||||||||||||            |||         :::::||||||||||:::
Db    142 GAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGTGTGCAGA 201

Qy     88 GlnLeuGlyCysProPheSerPheAlaMet---------PheArgPheGlyGlnAlaVal 104
          ||||||||||||         :::                      ||||||      :::
Db    202 CAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTGGGCCAGGACTT 261

Qy    105 ThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeu 124
                   |||   |||||||||         ||||||   |||   ||||||   :::
Db    262 ---------GGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTC 312

Qy    125 TrpGluCysGlnHis---------ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141
          :::|||:::|||            :::::::::   :::            ||||||
Db    313 AGTGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGG 372

Qy    142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161
```

```
Db     373 GATGCTGGAGTAGTCTGCTCAGGA-----------TTTGTGCGTCTGGCTGGAGGGGAT 420

Qy     162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
Db     421 GGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGAT 480

Qy     182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
Db     481 GGGAACTTCACACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCT 540

Qy     202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
Db     541 GTGTCTGTCCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAG 600

Qy     222 IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241
Db     601 TTCAGGTGTGAGGGGGAGGAGCCTGAGCTCTGGGTCTGCCCAGAGTGCCCTGTCCAGGG 660

Qy     242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu 261
Db     661 GGCACGTGTCACCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGAAGTCCGG 720

Qy     262 LeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281
Db     721 CTC---ATGACAAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGA 777

Qy     282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLys 301
Db     778 CAATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGT 837

Qy     302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321
Db     838 CAGCTCGGCTGTGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGT 897

Qy     322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
Db     898 GATCAGATCCTAACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGT 957

Qy     342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
Db     958 CCTGTGACTGCCCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGC 1017

Qy     362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
Db    1018 TCAGGAAAC--------------------------------------------------- 1026

Qy     382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
Db    1027 ---CAGATCCAGGTGCTTCCCCAG------------------------------------ 1047

Qy     402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
Db    1048 -----------------------------TGCAACGACTCCGTG--------------- 1062

Qy     422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
Db    1063 TCTCAACCTACAGGCTCTGCGGCCTCA-------------------------------- 1089

Qy     442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461
Db    1089 ------------------------------------------------------------ 1089

Qy     462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
Db    1090 GAGGACAGCGCCCCCTACTGCTCAGACAGCAGG------CAGCTCCGCCTGGTGGACGGG 1143

Qy     482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
Db    1144 GGCGGTCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGT 1203

Qy     502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
Db    1204 GATGACGGCTGGGACCTGGACGATGCCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAA 1263

Qy     522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
Db    1264 GCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGAC 1323

Qy     542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
Db    1324 AACTTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGG 1383

Qy     562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
Db    1384 CAGCACAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCA------GAGTTCCTG 1437

Qy     582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601
Db    1438 GCCCTCAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAAT 1497

Qy     602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCys 621
Db    1498 GGGACCTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGC 1557

Qy     622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
Db    1558 AGACAGCTTGGCTGTGGGGACAGT------------GGAACCCTCAACTCTTCTGTTGCT 1605

Qy     641 ---------GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
Db    1606 CTTAGAGAAGGTTTTAGGCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACC 1665

Qy     658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
Db    1666 TCTCTCTGGCAGTGTCCTTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAA 1725

Qy     678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697
Db    1726 GCCTATATCTGGTGTGCAGACAGC------AGACAGATCCGCCTGGTGGATGGAGGTGGT 1779

Qy     698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
Db    1780 CGCTGCTCTGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGAC 1839

Qy     718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIle 737
Db    1840 CGCTGGGACCTGGACGATGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTG 1899

Qy     738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
Db    1900 GACGCCACTGTCTCTTCCTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTG 1959

Qy     758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
Db    1960 AACTGCAGAGGAGAGGAGTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACAC 2019

Qy     778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
Db    2020 AACTGCAATCATCAAGAAGATGCAGGAGTCATCTGCTCAGGATTT------GTGCGTCTG 2073

Qy     798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
Db    2074 GCTGGAGGAGATGGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACC 2133

Qy     818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
Db    2134 CCAGTGTCTGATGGAAACTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGA 2193

Qy     838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
Db    2194 TGTGGCAAGGCTGTGTCTGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTC 2253

Qy     858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
Db    2254 TGGGCTGAAGAGTTCAGGTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTG 2313
```

```
Qy    878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
Db   2314 CCCTGTCCAGGAGGCACATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTAC 2373

Qy    898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
Db   2374 ACAGAAGTCCAGCTTATGAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAG 2433

Qy    917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
Db   2434 ATCTCTGGACGATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTT 2493

Qy    937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
Db   2494 GTCTGTCGTCAGCTCGGCTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTG 2553

Qy    957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
Db   2554 GAAGGAGGTGATCAGATCTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTG 2613

Qy    977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
Db   2614 TGGAGTTGTCCTGTGACTGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCT 2673

Qy    997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
Db   2674 GTGATCTGCTCAGGAAACCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAA 2733

Qy   1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
Db   2734 CCTGCAGGCTCTGCGGCCTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTC 2793

Qy   1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
Db   2794 CGCCTGGTGGACGGGGGCGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCC 2853

Qy   1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
Db   2854 TGGGGCACCATCTGTGATGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAG 2913

Qy   1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
Db   2914 CTGGGCTGTGGAGAAGCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGG 2973

Qy   1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
Db   2974 CCCATCTGGCTGGACGACCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCT 3033

Qy   1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
Db   3034 TCCCGGGGCTGGGGGCGGCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCA 3093

Qy   1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
Db   3094 GAGTTCCTGGCCCTCAGGATGGTGAGC------GAGGACCAGCAGTGTGCTGGGTGGCTG 3147

Qy   1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
Db   3148 GAGGTTTTCTACAACGGGACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACT 3207

Qy   1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
Db   3208 GTGTCCGTGATCTGCAGACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTT 3267

Qy   1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
Db   3268 GGTCTCAGGGAAGGTTCTAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGAT 3327

Qy   1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
Db   3328 ACCTCTCTCTGGCAGTGTCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAG 3387

Qy   1237 GluThrTrpIleThrCysGlu------------------------------------- 1243
Db   3388 GAAGCCTACATCTCATGTGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGC 3447

Qy   1244 ---------AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
Db   3448 ACAGACAGAGAGAAGCTCCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAG 3507

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
Db   3508 GTGTGGCACAACGGCTCCTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCT 3567

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
Db   3568 GAGGTGGTGTGTCAGCAGCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCA 3627

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
Db   3628 TTTGGCCCTGGAAATGGGAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCC 3687

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
Db   3688 TCCCTGTGGGACTGTGTTGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGAT 3747

Qy   1341 AlaGlyValArgCysSerGly------------------------------------- 1347
Db   3748 GCTGGTGTGAGGTGCTCTGGTGTAAGGACAACATTGCCCACGACCACAGCAGGGACCAGA 3807

Qy   1348 ---------GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeu 1364
Db   3808 ACAACCTCAAATTCTCTCCCTGGCATCTTCTCCCTGCCTGGGGTTCTCTGCCTTATCCTG 3867

Qy   1365 SerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg 1384
Db   3868 GGGTCGCTTCTCTTCCTGGTCCTCGTCATCCTGGTGACTCAGCTACTCAGATGG---AGA 3924

Qy   1385 ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu 1404
Db   3925 GCAGAGCGC------------------AGAGCCTTATCCAGCTATGAAGATGCTCTTGCT 3966

Qy   1405 GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg 1424
Db   3967 GAAGCTGTGTATGAGGAGCTCGATTACCTTCTGACACAGAAGGAAGGTCTGGGCAGCCCA 4026

Qy   1425 ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu 1443
Db   4027 GATCAGATGACTGATGTCCCTGATGAAAATTATGATGATGCTGAAGAAGTACCAGTGCCT 4086

Qy   1444 GlyValLeuProAlaSerGluAlaThrLys 1453
Db   4087 GGAACTCCTTCTCCCTCTCAGGGGAATGAG 4116
```

RESULT 7
US-09-917-800A-1712
; Sequence 1712, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11

```
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1712
;   LENGTH: 4344
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_022849
US-09-917-800A-1712

Alignment Scores:
Pred. No.:              1.78e-105       Length:         4344
Score:                  1142.50         Matches:        331
Percent Similarity:     41.73%          Conservative:   136
Best Local Similarity:  29.58%          Mismatches:     343
Query Match:            14.04%          Indels:         312
DB:                     10              Gaps:           39

US-09-759-130B-381 (1-1453) x US-09-917-800A-1712 (1-4344)

Qy    258 SerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeu 277
           |||   |||   :::|||||||||   |||   :::||||||   |||||||||||:::
Db     10 TCTGGTTTGGCTGTGAGGCTGGTGAATGGAGGAGACAGGTGTCGGGGTCGCGTGGAGATC 69

Qy    278 LysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAsp 297
                   ||||||   ||||||||:::|||         |||:::        |||:::
Db     70 CTTTACCAGGGTTCCTGGGGCACCATGTGTGATGACAGCTGGGACATCAATGATGCCAAC 129

Qy    298 ValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeu 317
           |||||||||:::|||||||||||||||   ||||||         |||
Db    130 GTGGTGTGCAGGCAGCTGGGCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAGTTT 189

Qy    318 GlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPhe 337
                   |||         :::   ||||||   |||:::|||   |||:::|||::::::
Db    190 GGACAGGGTCTGGGTCCCATTGTTCTGGATGACGTGGCCTGTAGAGGACATGAGGCCTAT 249

Qy    338 LeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspVal 357
           ||||||   |||   |||   |||   ::::::   :::|||   ||||||   |||
Db    250 CTGTGGAGCTGCTCCCACCGAGGCTGGCTCTCTCATAACTGTGGACATCAGGAGGATGCT 309

Qy    358 SerValIleCysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsn 377
              |||||||||||||||                     :::      |||:::
Db    310 GGAGTGATCTGCTCAGAT-----------------------TCTCAAACAAGCAGTCCC 345

Qy    378 CysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsn 397
                   |||                        ||||||      :::
Db    346 ACACCCGGT-------------------------TGGTGG---------AACCCCGGG 369

Qy    398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417
                   |||         :::         ::::::
Db    370 GGCACAAATAACGATGTGATCTACGACACTCAAGAAACC--------------------- 408

Qy    418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437
              :::      ::::::   |||:::         |||   |||
Db    409 ---ACAGAAACTTCTCAAACAAGCAGTCCCACACCTGATTGGTGG--------------- 450

Qy    438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr 457
                 :::      |||      ||||||               :::
Db    451 ------AACCATGGGGGCACAATTAATGATGTGATCTATGAC---------ACTCAAGAA 495

Qy    458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477
                   :::|||:::|||                        |||   :::|||
Db    496 ACCACAGAAGGAACAGATTCTGGT-----------------------TTGGCTGTGAGG 531

Qy    478 LeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrp 497
           ||||||            |||   ||||||:::|||:::   |||||||||   |||
Db    532 CTGGTGAATGGAGGAGACAGGTGTCGGGGTCGTGTGGAGATCCTTTACCAGGGTTCCTGG 591

Qy    498 GlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeu 517
           ||||||||||||   |||   |||         :::|||   |||||||||:::||||||
Db    592 GGCACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAACGTGGTGTGCAGGCAGCTG 651

Qy    518 GlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyPro 537
           |||||||||      :::         |||         |||   :::   ||||||
Db    652 GGCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAGTTTGGACAGGGCTCTGGGTCC 711

Qy    538 IleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHis 557
           |||   ||||||||||||:::|||   |||:::|||:::   :::|||   |||   |||
Db    712 ATTGTTCTGGATGACGTGGCCTGTAGAGGACATGAGGCCTATCTGTGGAGCTGCTCCCAC 771

Qy    558 SerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer--- 576
              ||||||      |||||||||   |||:::||||||         |||   ||||||
Db    772 CGAGGCTGGCTCTCTCATAACTGTGGACATCAGGAGGATGCTGGAGTCATCTGTTCATAT 831

Qy    576 ------------------------------------------------------------ 576

Db    832 TCTCAAACAAGCAGTCCCACACCCGATTCTCAAACAAGCAGTCCCACACCCGGTTGGTGG 891

Qy    577 -----------------GlyAspAlaThrTrpGly----------------------- 582
                               |||   ::::::|||
Db    892 AACCCCGGGGGCACAAATAACGATGTGTCCTATGGACCCGAACAGACCACAGACGCAACA 951

Qy    583 --------------LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGlu 597
                        :::|||||||||   |||   :::||||||   ||||||:::|||
Db    952 GATTCTGGTTTGGCTGTGAGGCTGGTGAATGGAGGAGACAGGTGTCAGGGTCGTGTGGAG 1011

Qy    598 ValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAla 617
          :::   :::||||||   ||||||||||||||||||||   |||:::::::|||   |||
Db   1012 ATCCTTTACCAGGGTTCCTGGGGTACCGTGTGTGACGACAGCTGGGACACCAAGGATGCC 1071

Qy    618 AlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsn 637
              |||||||||   ||||||   |||      ::::::         |||      :::
Db   1072 AACGTGGTGTGCAGGCAGCTGGTCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAC 1131

Qy    638 AlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
                |||   |||   |||   ||||||||||||:::|||   |||   |||:::
Db   1132 TTTGGACAAGGCTCTGGATCCATTGTTCTGGATGACGTGGCCTGTACAGGACATGAGGCC 1191

Qy    658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
             ||||||||||||   :::   ||||||   :::::::::|||   |||   ||||||
Db   1192 TATCTGTGGAGCTGCTCCCACCGAGGCTGGCTCTCTCATAACTGTGGCCACCATGAGGAT 1251

Qy    678 ValGlyValIleCysSerAspAlaSer--------------------AspMet------ 688
             |||||||||||||||||||||                        ||||||
Db   1252 GCTGGAGTCATCTGTTCAGATGCCCAAACCCAGAGCACAACCTGGCCAGATATGTGGCCT 1311

Qy    689 ------------Glu-LeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValGl 704
                      |||  |||:::   :::||||||   ::::::   :::      |||
Db   1312 ACTACCACTCCAGAAACTACAACAGATTGGTGGACT-ACAAAATATTCTTCCTCTGTTCC 1370

Qy    704 uValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAl 724
                   :::                     |||:::   |||
Db   1371 TACAACACAATTCCCCACCATA------------GCCGATTGGTGGACAACTCCTTCTCC 1418

Qy    724 aGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer----------- 740
            |||                     ||||||      :::   :::
Db   1419 GGAATACACC---------------TGTGGAGGTTTACTGACCCTACCCTATGGGCAGTT 1463

Qy    741 -ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysTh 760
             |||::::::                             :::|||:::   |||
Db   1464 TTCCAGCCCATACTACCCTGGAAGC------------TATCCTAACAATGCCAGATGT-- 1509

Qy    760 rGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHi 780
                            ||||||
```

```
Db   1510 ---------------TTGTGG------------------------------------ 1515

Qy    780 sLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAl 800
                           :::      |||:::     :::     :::|||
Db   1516 ---------------AAAATTTTCGTCTCCAGCATGAACCGTGTGACAGTGGTCTTCAC 1559

Qy    800 aAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCy 820
            |||:::          |||                                   ||
Db   1560 AGATGTGCAGCTTGAAGGAGGT----------------------------------TG 1583

Qy    820 sAspSerAspPheSerLeu-----------------HisAlaAlaAsnValLeuCysAr 834
          |:::  |||:::  |||                   :::  :::::::::::    ||
Db   1584 CAACTATGACTACATCCTGGTTTTTGATGGCCCTGAAAACAATTCTTCTCTCATTGCTCG 1643

Qy    834 gGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly-- 853
          |                                 |||  |||  |||
Db   1644 G--------------------------------GTTTGTGATGGGTTCAATGGATCTTT 1670

Qy    854 --------------------------------------AsnGlyLeuThrTrpAlaGl 860
                                                :::|||
Db   1671 CACCTCAACCCAGAACTTCATGTCTGTAGTCTTTATCACGGATGGCAGTGTCACGAGGAG 1730

Qy    860 uLysPheGlnCysGlu--------------------GlySerGluThrHisLeuAlaLe 873
                ||||||   :::                        |||   |||
Db   1731 AGGGTTCCAAGCTGACTACTACTCCACTCCTATCAGCACCAGCACAACCTCTCCAACGAC 1790

Qy    873 uCysProIleVal-----------------GlnHisProGluAspThrCys------Il 885
              |||||||||                 ||||||  ||||||          ::
Db   1791 GTTCCCGATCGTTACTGATTGGTGGACAACTCCTTCTCCGGAATACACCTGTGGAGGTTT 1850

Qy    885 eHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGl 905
          :   :::         |||         |||  |||
Db   1851 ACTGACCCTACCCTATGGGCAGTTTTCCAGCCCATAC----------------------- 1887

Qy    905 yLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCy 925
                                                    :::   ||||||
Db   1888 ------------------------------------------TACCCTGGAAGCTAT-- 1902

Qy    925 sAspThrHisTrpAspProGluAspAlaArgValLeuCysArg---------------- 939
                           |||   :::||||||  |||   :::
Db   1903 ----------------CCTAACAATGCCAGATGTTTGTGGAAAATTTTCGTCCCCAGCAT 1946

Qy    940 ------------------------------GlnLeuSerCysGlyThr----------- 945
                                        ||||||    |||
Db   1947 GAACCGTGTGACAGTGGTCTTCACAGATGTGCAGCTTGAAGGAGGTTGCAACTATGACTA 2006

Qy    946 -AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHi 965
             |||      |||   :::|||                ||||||
Db   2007 CATCCTGGGTTTTGATGGTCCTGAATACAATTCTTCTCTCATTGCTCGGGTT-------- 2058

Qy    965 sArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGl 985
                   |||   |||:::                 :::   |||
Db   2059 ---------TGTGATGGGTCCAATGGATCTTTCACCTCAACCCAG------------- 2094

Qy    985 yAlaProProCysIleHisGlyAsnThrValSerVal------IleCysThrGlySerLe 1003
                              |||   :::||||||     |||     ||||||::
Db   2095 ---------------------AACTTCATGTCTGTAGTCTTTATCACGGATGGCAGTGT 2132

Qy   1003 uThrGlnProLeuPhe--------------------ProCysLeuAlaAsnValSerAs 1016
          :|||:::      |||                     |||          :::  :::
Db   2133 CACGAGGAGAGGGTTCCAAGCTGACTACTACTCCACTCCTATCAGGACCAGCACAACTCC 2192

Qy   1016 pProTyrLeuSerAlaValProGluGly------SerAlaLeuIleCysLeuGluAspLy 1034
            |||         :::        |||       |||:::|||:::
Db   2193 TCCAACGACGTTCCCGATCATTACTGGAAATGATTCTTCATTGGTG------------- 2238

Qy   1034 sArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAs 1054
                ||||||||||||:::|||    :::||||||   ||||||||||||||     :::
Db   2239 ----CTGAGGCTGGTAAATGGAACAAACCGGTGTGAGGGCCGAGTGGAGATCTTGTACAG 2294

Qy   1054 pGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCy 1074
            |||    |||             ||||||   ||||||:::::|||||:::||||||||
Db   2295 AGGCTCTTGGGTACCGTGTGCCGACGACAGCTGGGACATCAATGATGCCAATGTGGTCTG 2354

Qy   1074 sGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGl 1094
          |::::::|||||||||||||||     |||    :::|||        :::|||    |||||||:::||
Db   2355 CAGACAGCTCGGTTGTGGCTCTGCTCTGTCTGCTCCAGGAAATGCTTGGTTTGGTCAGGG 2414

Qy   1094 ySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGl 1114
          ||||||     |||     |||||||||::::::|||:::|||     ||||||||||||||
Db   2415 TTCAGGGCTCATTGTCCTGGATGATGTGTCTTGCTCTGGGTATGAGTCCCACCTGTGGAA 2474

Qy   1114 nCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIl 1134
            |||                ||||||        |||:::|||||||||    |||||||||||||||
Db   2475 TTGTCGTCACCCTGGCTGGCTTGTTCATAATTGTCGTCATGTTGAGGATGCAGGAGTCAT 2534

Qy   1134 eCysSer------GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCy 1152
          ||||||           :::    |||    |||         |||         ::::::     ||
Db   2535 TTGCTCACTCCCTGATCCGACTCC-CTCTCCTGGTCCAGTTTGGACAAGTCCTCCTTTTG 2593

Qy   1152 sAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIl 1172
          |          :::|||      ::::::                 |||    |||:::|||
Db   2594 T------AAACTATACTTGTGGAGGTTTCCTGACTGGAC-TCTCTGGGCAATTTTCTAGC 2646

Qy   1172 eThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValVa 1192
                |||     :::          |||:::                   ||||||
Db   2647 CCATACTACCCTGGGAGCTATCCTAATAATGCCAGATGTTTGTGGA------------- 2692

Qy   1192 lSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp-------- 1209
                             |||:::|||     :::      :::|||          :::
Db   2693 -------ACATTGAAGTCCCAAACAACTACCGCGTGACTGTGGTCTTCAGAGATGTGCAG 2745

Qy   1210 -----------------------------------IleGlnCysProLysThr---Hi 1216
                                             :::      |||   |||      ||
Db   2746 CTGGAAGGGGGCTGCAACTATGACTATATAGAGATTTTTGATGGCCCCCACCACAGTTCA 2805

Qy   1216 sIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro 1234
          |:::|||:::                :::      ||||||         :::         |||
Db   2806 CCTCTCATTGCCCGGGTTTGTGATGGGGCCATGGGCTCTTTCACTTCAACATCCA 2860
```

RESULT 8
US-09-147-947-5
; Sequence 5, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoko
; APPLICANT: YAMAGUCHI, No. US20020160490A1omi
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 2562
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   OTHER INFORMATION:
US-09-147-947-5

Alignment Scores:

Pred. No.:	4.26e-74	Length:	2562
Score:	830.00	Matches:	260
Percent Similarity:	40.26%	Conservative:	106
Best Local Similarity:	28.60%	Mismatches:	349
Query Match:	10.20%	Indels:	195
DB:	9	Gaps:	34

US-09-759-130B-381 (1-1453) x US-09-147-947-5 (1-2562)

```
Qy      132 TrpGlySerHisAsnCysTyrHisGlyGlu-----------------AspValGlyVal 145
            ||||||         |||     ||||||                 |||   |||
Db      112 TGGGGC---------TGCCCCGCCGGCGAGCCATGGGTCAGCGTGACGGACTTCGGCGCC 162

Qy      146 AsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAsp-GlyAsnAsnSerCysSe 165
               |||                   ||||||    :::    :::  |||   ||
Db      163 CCGTGT--------------------CTGCGGTGGGCGGAGGTGCCACCCTTCCTGGAG 201

Qy      165 rGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAs 185
            |||||||                  :::|||   |||     |||:::|||
Db      202 CGGTCGC---------------CCCCAGCGAGCTGGGCTCAGCTGCGAGGACAGC----- 241

Qy      185 nLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSe 205
                                               :::|||   |||:::::
Db      242 ------------------GCCACAACTTTTGTCGGAGCCCCGACGGCGCGGGCAGACCC 282

Qy      205 rGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAspIleLeuCysGl 225
            ||||||               |||   |||
Db      283 TGGTGTTTCTACGGAGACGCCCGTGGCAAGG----------------------------- 313

Qy      225 nGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSe 245
                               |||                ||||||     ||||||
Db      314 ------------------TGGAC----------------TGGGGCTACTGCGACTGCAG 338

Qy      245 rHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuVa 265
             |||                                            :::||||||
Db      339 ACACGGA--------------------------------------TCAGTACGACTTCG 359

Qy      265 lGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyTh 285
              ||||||   |||          |||   ||||||:::         |||   ||||||||
Db      360 TGGCGGCAAAAATGAGTTTGAAGGCACAGTGGAAGTATATGCAAGTGGAGTTTGGGGCAC 419

Qy      285 rValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeu----- 303
             |||||||            |||:::::::::   |||   |||:::|||   ||||||
Db      420 TGTCTGTAGCAGCCACTGGGATGATTCTGATGCATCAGTCATTTGTCACCAGCTGCAGCT 479

Qy      304 ---------GlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGl 320
                      |||              |||:::||||||                ||
Db      480 GGAGGAAAAGGAATAGCAAAACAAACCCCGTTTTCTGGACTG---------------GG 524

Qy      320 ySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAs 340
            |          ::::::         |||   |||   |||:::|||      :::
Db      525 CCTTATTCCCATTTATTGGAGCAATGTCCGTTGCCGAGGAGATGAAGAAAATATACTGCT 584

Qy      340 pCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIl 360
              |||                        |||      :::         :::|||
Db      585 TTGTGAAAAAGACATCTGGCAGGGTGGGGTGTGTCCTCAGAAGATGGCAGCTGCTGTCAC 644

Qy      360 eCys------SerAspGlyAlaAspLeuGlu---LeuArgLeuAlaAspGlySerAsnAs 377
              |||      |||   |||               :::|||||||||   ||||||:::
Db      645 GTGTAGCTTTTCCCATGGCCCAACGTTCCCCATCATTCGCCTTGCTGGAGGCAGCAGTGT 704

Qy      377 nCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAs 397
                     ||||||||||||:::               ||||||   |||:::||||||
Db      705 GCATGAAGGCCGGGTGGAGCTCTACCATGCTGGCCAGTGGGGAACCGTTTGTGATGACCA 764

Qy      397 nTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPh 417
              |||   :::         |||   |||:::|||:::|||||||||
Db      765 ATGGGATGATGCCGATGCAGAAGTGATCTGCAGGCAGCTGGGCCTCAGTGGCATTGCCAA 824

Qy      417 eGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSe 437
                     :::|||         :::      :::   ::::::   :::
Db      825 AGCATGGCATCAGGCATATTTGGGGAAGGGTCTGGCCCAGTTATGTTGGATGAAGTACG 884

Qy      437 rCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgTh 457
             |||||||||||||||      ::::::      |||               :::
Db      885 CTGCACTGGGAATGAGCTTTCAATTGAGCAGTGTCCAAAGAGCTCCTGGGGAGAGCATAA 944

Qy      457 rCysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuAr 477
              |||      :::   ||||||||||||   |||:::            |||      :::||
Db      945 CTGTGGCCATAAAGAAGATGCTGGAGTGTCCTGTACCCCTCTAACAGATGGGGTCATCAG 1004

Qy      477 gLeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTr 497
             ||||   |||                    ||||||||||||||    |||:::|||:::||
Db     1005 ACTTGCAGGTGGGAAAGGCAGCCATGAGGGTCGCTTGGAGGTATATTACAGAGGCCAGTG 1064

Qy      497 pGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLe 517
             ||||||||||||   |||   |||:::         |||   ||||||||:::|||||
Db     1065 GGGAACTGTCTGTGATGATGGCTGGACTGAGCTGAATACATACGTGGTTTGTCGACAGTT 1124

Qy      517 uGly------CysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSe 535
             ||||       ||||||                          :::|||:::|||:::||
Db     1125 GGGATTTAAATATGGTAAACAAGCA---------TCTGCCAACCATTTTGAAGAAAGCAC 1175

Qy      535 rGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCy 555
            :|||||||||||||||||||||||||||||   |||   |||:::          ||
Db     1176 AGGGCCCATATGGTTGGATGACGTCAGCTGCTCAGGAAAGGAAACCAGATTTCTTCAGTG 1235

Qy      555 sGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCy 575
            |               ||||||:::|||:::|||   ||||||||||||||   :::   ||
Db     1236 TTCCAGGCGACAGTGGGGAAGGCATGACTGCAGCCACCGCGAAGATGTTAGCATTGCCTG 1295

Qy      575 s------SerGlyAspAla--------------ThrTrpGlyLeuArgLeuValGlyGl 588
            |          |||:::               :::   :::||||||:::   ||
Db     1296 CTACCCTGGCGGCGAGGGACACAGGCTCTCTCTGGGTTTTCCTGTCAGACTGATGGATGG 1355

Qy      588 ySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrValCy 608
            |    |||:::          ||||||:::||||||:::   |||:::|||||||||:::||
Db     1356 AGAAAATAAGAAAGAAGGACGAGTGGAGGTTTTTATCAATGGCCAGTGGGGAACAATCTG 1415

Qy      608 sAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAspCysProSe 628
            ||||||||||||         |||      |||||||||:::|||      ||||||
Db     1416 TGATGATGGATGGACTGATAAGGATGCAGCTGTGATCTGTCGTCAGCTTGGCTACAAGGG 1475

Qy      628 rSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAs 648
                             :::                  |||   |||   |||   :::||
Db     1476 TCCTGCCAGAGCAAGAACCATGGCTTACTTTGGAGAAGGAAAAGGACCCATCCATGTGGA 1535

Qy      648 pAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGl 668
            |:::|||   |||   |||:::|||         |||      |||              ||
Db     1536 TAATGTGAAGTGCACAGGAAATGAGAGGTCCTTGGCTGACTGTATCAAGCAAGATATTGG 1595

Qy      668 yAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAlaSerAspMe 688
            |   ::::::|||   ||||||||||||   ||||||||||||            |||
Db     1596 AAGACACAACTGCCGCCACAGTGAAGATGCAGGAGTTATTTGT-----------GATTA 1643

Qy      688 tGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValGluValAsnValGl 708
               :::          |||:::|||                    |||   ::
Db     1644 TTTTGGCAAGAAGGCCTCAGGTAACAGT-----------------------AATAAAGA 1679

Qy      708 nGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluValValCy 728
            :         :::
Db     1680 GTCCCTCTCATCTGTT---------------------------------------- 1695

Qy      728 sArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluProHisPheThrGluAr 748
                          ||||||       :::                          ||
Db     1696 -------------TGTGGCTTGAGATTA--------------------CTGCACCGTCG 1721

Qy      748 gThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAs 768
            |         |||:::         :::         ||||||
Db     1722 GCAGAAGCGGATCATTGGTGGGAAAAATTCTTTAAGGGGTGGT---------------- 1764

Qy      768 pCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIl 788
                        |||   |||   |||   :::         |||
Db     1765 ----------TGGCCTTGG---CAGGTTTCCCTCCGGCTGAAG---------------- 1794

Qy      788 eCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMetProCysSerGlyArgVa 808
                   |||:::|||         ||||||:::                  |||
Db     1795 ----TCATCCCATGGAGATGGCAGGCTCCTC---------------TGCGGGGCT----- 1830

Qy      808 lGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAl 828
                                     ::::::         |||   :::   |||   ||
Db     1831 ---------------------------ACGCTCCTGAGTAGCTGCTGGGTCCTCACAGC 1862

Qy      828 aAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAs 848
            ||||:::          |||         |||         :::   ::::::   ||||||||
```

```
Db    1863 AGCACACTGTTTCAAGAGGTATGGCAACAGCACTAGGAGCTATGCTGTTAGGGTTGGAGA 1922

Qy     848 pHis--------------------PheGlyLysGlyAsnGly---------LeuThrTr 858
           |:::                    |||   :::   |||          :::
Db    1923 TTATCATACTCTGGTACCAGAGGAGTTTGAGGAAGAAATTGGAGTTCAACAGATTGTGAT 1982

Qy     858 pAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGl 878
               :::::::::   :::   |||:::      :::||||||      :::    ||
Db    1983 TCATCGGGAGTATCGACCCGACCGCAGTGATTATGACATAGCCCTGGTTAGATTA---CA 2039

Qy     878 nHisProGluAspThrCysIleHis---SerArgGluValGlyValValCysSerArgTy 897
           |   ||||||:::   |||         |||      |||         |||
Db    2040 AGGACCAGAAGAGCAATGTGCCAGATTCAGCAGCCATGTTTTGCCAGCCTGTTTACCACT 2099

Qy     897 rThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnVa 917
                       |||            |||   |||
Db    2100 CTGGAGAGAGAGGCCACAGAAAACAGCATCCAACTGT--------------------TA 2138

Qy     917 lLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLe 937
            :::      ||||||                             :::      |||
Db    2139 CATAACAGGATGGGGT--------------------------GACACAGGACGAGCCTA 2171

Qy     937 uCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGl 957
              |||   |||            |||:::            ::::::      ||
Db    2172 TTCAAGAACACTACAA---CAAGCAGCCATTCCCTTACTTCCTAAAAGGTTTTGTGAAGA 2228

Qy     957 uArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsn---------GluSe 974
           ||||           :::         |||      |||   ||||||
Db    2229 ACGTTATAAGGGTCGGTTTACAGGGAGAATGCTTTGTGCTGGAAACCTCCATGAACACAA 2288

Qy     974 rLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaPro-ProCysIleHisGlyAsnT 994
              :::|||:::||||||             |||   |||   |||::::::|||
Db    2289 ACGCGTGGACAGCTGCCAGGGAGAC---AGCGGAGGACCACTCATGTGTGAACGGCCCGG 2345

Qy     994 hrValSerValIleCysThrGly 1001
              :::      |||   |||
Db    2346 AGAGAGCTGGGTGGTGTATGGGG 2368
```

```
RESULT 9
US-09-918-995-32438        20030073623
; Sequence 32438, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32438
;   LENGTH: 473
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(473)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32438

Alignment Scores:
Pred. No.:              6.13e-74      Length:        473
Score:                  818.00        Matches:       143
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            10.05%        Indels:        0
DB:                     9             Gaps:          0

US-09-759-130B-381 (1-1453) x US-09-918-995-32438 (1-473)
```

```
Qy    1194 LeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysPro 1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 CTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCT 101

Qy    1214 LysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSer 1233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 AAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGC 161

Qy    1234 ProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThr 1253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 CCAGCAGAAGAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACC 221

Qy    1254 GluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAsp 1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 GAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGAC 281

Qy    1274 SerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeu 1293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 TCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTG 341

Qy    1294 AlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMet 1313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATG 401

Qy    1314 ArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSer 1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 CGGTGCAAAGGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGT 461

Qy    1334 AspCysGly 1336
           |||||||||
Db     462 GACTGTGGA 470
```

```
RESULT 10
US-09-147-947-3
; Sequence 3, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
;  APPLICANT: TSURUOKA, No. US20020160490A1uo
;  APPLICANT: YAMASHIRO, Kyoko
;  APPLICANT: YAMAGUCHI, No. US20020160490A1omi
;  TITLE OF INVENTION: No. US20020160490A1el Serine Protease
;  FILE REFERENCE: 001560-349
;  CURRENT APPLICATION NUMBER: US/09/147,947A
;  CURRENT FILING DATE: 1997-03-24
;  EARLIER APPLICATION NUMBER: PCT/JP98/03324
;  EARLIER FILING DATE: 1998-07-24
;  EARLIER APPLICATION NUMBER: JP 9/213969
;  EARLIER FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2614
;   TYPE: DNA
;   ORGANISM: Mouse
;   FEATURE:
;   OTHER INFORMATION:
US-09-147-947-3

Alignment Scores:
Pred. No.:              3.99e-63      Length:        2614
Score:                  722.00        Matches:       207
Percent Similarity:     41.54%        Conservative:  85
Best Local Similarity:  29.45%        Mismatches:    260
Query Match:            8.87%         Indels:        153
DB:                     9             Gaps:          25

US-09-759-130B-381 (1-1453) x US-09-147-947-3 (1-2614)

Qy      10 IleAspPheGlyArgCysCysCysHisGlnAsnLeuPheSerAlaValValThrCysIle 29
           :::|||:::|||   |||   |||   |||               |||:::
Db     692 GTAGACTGGGGCTACTGCGATTGTGGTCAAGGCCCGGCGTTGCCCGTCATT--------- 742

Qy      30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGluLeuArg 49
```

```
                                                                  |||
Db     743 ----------------------------------------------------------CGC 745

Qy      50 LeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrp 69
           |||||| . |||:::                 |||   ||||||:::         |||||||||
Db     746 CTTGTTGGTGGGAACAGTGGGCATGAAGGTCGAGTGGAGCTGTACCACGCTGGCCAGTGG 805

Qy      70 GlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeu 89
           |||||:::|||||||||   |||:::          :::  .|||:::|||:::||||||
Db     806 GGGACCATCTGTGACGACCAATGGGACAATGCAGACGCAGACGTCATCTGTAGGCAGCTG 865

Qy      90 ---------GlyCysProPheSerPheAlaMetPheArgPheGlyGlnAlaValThrArg 106
                    |||          ::::::          |||||||:::
Db     866 GGGCTCAGTGGCATTGCCAAAGCATGGCATCAGGCACATTTTGGGGAAGGA--------- 916

Qy     107 HisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGlu 126
              |||   |||   ||||||:::|||   |||   ||||||||   ::::::   :::
Db     917 TCTGGCCCAATATTGTTGGATGAAGTACGCTGCACCGGAAACGAGCTGTCAATTGAGCAA 976

Qy     127 CysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsn 146
           |||            ||||||   |||||||||   |||   ||||||   ||||||:::
Db     977 TGTCCAAAGAGTTCCTGGGGCGAACATAACTGTGGCCATAAAGAAGATGCTGGAGTGTCT 1036

Qy     147 CysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGly 166
           |||            :::   ::::||||||    |||   ::::::         |||
Db    1037 TGTGTTCCTCTAACAGATGGTGTCATCAGACTGGCAGGAGGAAAAAGTACCCATGAAGGT 1096

Qy     167 ArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeu 186
           |||:::||||||   ::::::   :::||||||||||:::||||||||||||||
Db    1097 CGCCTGGAGGTCTACTACAAGGGGCAGTGGGGGACAGTCTGTGATGATGGCTGGACTGAG 1156

Qy     187 AsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGly 206
                    |||   ||||||   ||||||                     ||||||
Db    1157 ATGAACACATACGTGGCTTGTCGACTGCTGGGATTTAAATACGGCAAACAGTCCTCT--- 1213

Qy     207 ValValAsnSerProAlaValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGly 226
           |||   :::         ||||||||||||||||||||:::         |||   |||
Db    1214 GTGAACCATTTTGATGGCAGCAACAGGCCCATATGGCTGGATGACGTCAGCTGCTCAGGA 1273

Qy     227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
             |||::::::      |||   |||   ||||||   ||||||||||||||
Db    1274 AAAGAAGTCAGCTTCATTCAGTGTTCCAGGAGACAGTGGGGAAGGCATGACTGCAGCCAT 1333

Qy     247 AsnGluAspValThrLeuThrCysTyrAspSerSerAsp--------------------- 259
             |||||||||   ||||||||||||      ||||||
Db    1334 AGAGAAGATGTGGGCCTCACCTGCTATCCTGACAGCGATGGACATAGGCTTTCTCCAGGT 1393

Qy     260 LeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIle 279
                 :::|||||||||   |||   |||:::      ||||||||||||:::   :::
Db    1394 TTTCCCATCAGACTAGTGGATGGAGAGAATAAGAAGGAAGGACGAGTGGAGGTTTTTGTC 1453

Qy     280 GlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValVal 299
              |||:::|||||||||:::|||           |||   :::      |||   |||:::
Db    1454 AATGGCCAATGGGGAACAATCTGCGATGACGGATGGACCGATAAGCATGCAGCTGTGATC 1513

Qy     300 CysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSer 319
           |||:::|||||||||                     :::   :::
Db    1514 TGCCGGCAGCTTGGCTATAAGGGTCCTGCCAGAGCAAGGACTATGGCTTATTTTGGGGAA 1573

Qy     320 GlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrp 339
           |||         :::   :::|||   |||  . |||:::|||||||||         |||
Db    1574 GGAAAAGGCCCCATCCACATGGATAATGTGAAGTGCACAGGAAATGAGAAGGCCCTGGCT 1633

Qy     340 AspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerVal 359
           ||||||                     :::|||   |||         |||      |||
Db    1634 GACTGTGTCAAACAAGACATTGGAAGGCACAACTGCCGCCACAGTGAGGATGCAGGAGTC 1693

Qy     360 IleCys--------------------SerAspGlyAlaAspLeuGluLeuArgLeuAla 372
           ||||||                     |||         :::   |||:::
Db    1694 ATCTGTGACTATTTAGAGAAGAAAGCATCAAGTAGTGGTAATAAAGAGATGCTC------ 1747

Qy     373 AspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGlu------GlnTrp 390
                  |||:::   |||      |||:::      |||      |||            |||

Db    1748 ------TCATCTGGATGTGGACTGAGGTTACTGCACCGTCG-GCAGAAACGGATCATTGG 1800

Qy     391 TrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu 410
           |||         :::|||   :::|||         ||||||
Db    1801 TGG---------GAACAATTCTTTAAGGGGTGCCTGGCCTTGGCA--------------- 1836

Qy     411 GlyCysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAsp 430
           |||   |||   :::         |||      |||   :::   :::
Db    1837 GGCTTCCCTCAGGCTGAGGTCGGCCCATGGAGACGGCAGGCTGCT--------------- 1881

Qy     431 IleTrp--------------------IleAsnSerIleSerCysThrGlyAsnGlu--- 442
           :::|||                     :::         |||||||||      :::
Db    1882 TTGTGGAGCTACCCTTCTGAGTAGCTGCTGGGTCCTGACAGCTGCACACTGCTTCAAAAG 1941

Qy     443 --------------------------SerAlaLeuTrpAspCysThrTyrAspGly--- 452
                                     |||:::   |||      :::|||   |||
Db    1942 GTACGGAAACAACTCGAGGAGCTATGCAGTTCGAGTTGGGGATTATCATACTCTGGTACC 2001

Qy     453 ------------------LysAlaLysArgThrCysPheArgArgSerAspAlaGly 465
                             :::               |||               |||:::
Db    2002 AGAGGAGTTTGAACAAGAAATAGGGGTTCAACAGATTGT-----------GATTCACAG 2049

Qy     466 ValIleCysSer-AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCy 485
             :::   :::  |||::::::|||   |||:::   ||||||               |||
Db    2050 GAACTACAGGCCAGACAGAAGCGACTATGACATTGCCCTGGTTAGATTGCAAGGACCA-- 2107

Qy     485 sTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTr 505
               |||                :::                  |||          ||
Db    2108 ----GGGGAGCAATGTGCCAGACTAAGCACCCACGTTTTGCCAGCCTGTTTACCTCTATG 2163

Qy     505 pSerThrArg------AsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
            |   |||.               |||      |||         |||   |||
Db    2164 GAGAGAGAGGCCACAGAAAACAGCCTCCAACTGTCACATAACAGGATGGGGAGACACAGG 2223

Qy     523 tHisValPheGlyMetThr-----------------------------------TyrPh 531
                 :::      |||                                     :::
Db    2224 TCGTGCCTACTCAAGAACTCTACAACAAGCTGCTGTGCCTCTGTTACCCAAGAGGTTTTG 2283

Qy     531 eLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAs 551
             ||||||               ::::::         |||   ||||||
Db    2284 TAAAGAGAGGTACAAGGGACTATTTACTGGGAGAATGCTCTGTGCTGGGAAC-------- 2335

Qy     551 nIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspVa 571
                                                  :::      ::::::
Db    2336 ------------------------------------------CTCCAAGAAGACAACCG 2352

Qy     571 lIleValThrCysSerGlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnAr 591
              :::   :::|||   ||||||:::                        ||||||
Db    2353 TGTGGACAGCTGCCAGGGAGACAGT--------------------GGAGGACCACTCAT 2391

Qy     591 gCysSerGlyArgLeuGlu---------ValTyrPheGlnGlyArgTrpGlyThrValCy 608
              |||         |||         ||||||            ||||||       ||
Db    2392 GTGTGAAAAGCCTGATGAGTCCTGGGTTGTGTATGGGGTGACTTCCTGGGGGTATGGATG 2451

Qy     608 s---------AspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAs 625
           |           |||         |||:::               :::      :::|||:::   |||||
Db    2452 TGGAGTCAAAGACAC-TCCTGGAGTTTATACCAGAGTCCCCGCCTTTGTACC---TTGGA 2507

Qy     625 p------CysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGl 643
           |       |||   |||:::                              |||||
Db    2508 TAAAAAGTGTCACCAGTCTGTAACT---------------------------TATGG 2537

Qy     643 yLysIle 645
            ||||:::
Db    2538 AAAGCTC 2544
```

RESULT 11
US-09-905-291A-147
; Sequence 147, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi

```
; APPLICANT:  Botstein, David
; APPLICANT:  Desnoyers, Luc
; APPLICANT:  Eaton, Dan L.
; APPLICANT:  Ferrara, Napoleone
; APPLICANT:  Filvaroff, Ellen
; APPLICANT:  Fong, Sherman
; APPLICANT:  Gao, Wei-Qiang
; APPLICANT:  Gerber, Hanspeter
; APPLICANT:  Gerritsen, Mary E.
; APPLICANT:  Goddard, A.
; APPLICANT:  Godowski, Paul J.
; APPLICANT:  Grimaldi, Christopher J.
; APPLICANT:  Gurney, Austin L.
; APPLICANT:  Hillan, Kenneth, J.
; APPLICANT:  Kljavin, Ivar J.
; APPLICANT:  Mather, Jennie P.
; APPLICANT:  Pan, James
; APPLICANT:  Paoni, Nicholas F.
; APPLICANT:  Roy, Margaret Ann
; APPLICANT:  Stewart, Timothy A.
; APPLICANT:  Tumas, Daniel
; APPLICANT:  Williams, P. Mickey
; APPLICANT:  Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION:  Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE:  2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 147
;   LENGTH: 1686
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-905-291A-147

Alignment Scores:
Pred. No.:            2.64e-60      Length:         1686
Score:                691.50        Matches:        148
Percent Similarity:   52.51%        Conservative:   40
```

```
Best Local Similarity:  41.34%        Mismatches:    139
Query Match:            8.50%         Indels:        31
DB:                     9             Gaps:          9

US-09-759-130B-381 (1-1453) x US-09-905-291A-147 (1-1686)

Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----------Phe 36
           |||||||||                   |||:::|||     |||          |||
Db     107 CTATTCTCC-----------------TTGATCCTTGCCATTTGCACCAGACCTGGATTC 148

Qy      37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
           |||   |||                              :::|||||||||   |||
Db     149 CTAGCGTCTCCATCTGGA--------------GTGCGGCTGGTGGGGGGCCTCCACCGC 193

Qy      57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
           |||   |||   ||||||||||:::   :::|||||||||||||||||||||||||||||
Db     194 TGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATGACGGC 253

Qy      77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
           |||:::                 |||:::|||::::::|||||||||      :::
Db     254 TGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313

Qy      97 Met---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
                           :::               :::     |||:::    :::
Db     314 ACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCAA 373

Qy     114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
              |||||||||   |||   |||         |||   :::|||:::      |||
Db     374 TCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAA------ 427

Qy     134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----------------- 147
              ::::::|||   |||   ||||||   |||   :::|||
Db     428 GTTTATGATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCT 487

Qy     148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
           :::                  |||:::||||||   ||||||         |||   ||||||
Db     488 TTCTCCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGC 547

Qy     168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
           ||||||||||||   |||   :::|||   |||:::|||      ||||||:::|||
Db     548 GTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGG 607

Qy     188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
              |||   ||||||||||||||||||||||||            :::   ::::::
Db     608 GCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGC 667

Qy     208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
              |||      |||      |||   ||||||||||||      :::   |||   |||
Db     668 TGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727

Qy     227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
              |||      |||   :::|||               ||||||   :::   |||:::|||
Db     728 CGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGAAGAACACCTGCAACCAT 787

Qy     247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
           :::||||||      :::   |||   |||         |||      ||||||||||||||
Db     788 GATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC------TTGAGACTAGTAGGA 841

Qy     267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
           |||   |||   |||   ||||||:::|||:::            :::|||   ||||||:::|||
Db     842 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGTC 901

Qy     287 CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
           |||               |||         ||||||||||||||||||||||||||||||
Db     902 TGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGGG 961

Qy     307 ThrAlaLeu------HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
              :::|||         |||                              |||      :::
Db     962 AAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCATC 1021

Qy     325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
           |||||||||   |||   |||||||||   |||      |||      |||:::|||
Db    1022 TGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAGA 1081
```

```
Qy    345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
                                 ||||||   ||||||   ||||||:::|||||||||||
Db   1082 TTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCA 1135

RESULT 12
US-09-902-853-147
; Sequence 147, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,853
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US/09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 147
;   LENGTH: 1686
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-902-853-147

Alignment Scores:
Pred. No.:              2.64e-60     Length:        1686
Score:                  691.50       Matches:       148
Percent Similarity:     52.51%       Conservative:  40
Best Local Similarity:  41.34%       Mismatches:    139
Query Match:            8.50%        Indels:        31
DB:                     9            Gaps:          9

US-09-759-130B-381 (1-1453) x US-09-902-853-147 (1-1686)

Qy     21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys------------Phe 36
          ||||||||                          |||:::|||      |||          |||
Db    107 CTATTCTCC-----------------TTGATCCTTGCCATTTGCACCAGACCTGGATTC 148

Qy     37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
          |||   |||                                :::|||||||||   |||
Db    149 CTAGCGTCTCCATCTGGA---------------GTGCGGCTGGTGGGGGGCCTCCACCGC 193

Qy     57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
          |||   |||   |||||||||:::   :::||||||||||||||||||||||||||||||
Db    194 TGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATGACGGC 253

Qy     77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
          |||:::                  |||:::|||:::::::||||||||          :::
Db    254 TGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313

Qy     97 Met---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
                           :::                 :::     |||:::   :::
Db    314 ACCCCTAGTGGTATTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCAA 373

Qy    114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
             |||||||||   |||   |||      |||   :::|||:::          |||
Db    374 TCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAA------ 427

Qy    134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------------------ 147
             :::::|||   |||   ||||||   |||   :::|||
Db    428 GTTTATGATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCT 487

Qy    148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
          :::                |||:::||||||   ||||||          |||   ||||||
Db    488 TTCTCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGC 547

Qy    168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
          ||||||||||||   |||   :::|||   |||:::|||       ||||||:::|||
Db    548 GTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGG 607

Qy    188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
             |||   ||||||||||||||||||||||||||||||      :::   :::::
Db    608 GCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGC 667

Qy    208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
             |||      |||      |||   ||||||||||||      :::   |||   |||
Db    668 TGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727

Qy    227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
             |||      |||   :::|||                  ||||||   :::   |||:::|||
Db    728 CGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCAT 787

Qy    247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
          :::||||||         :::   |||   |||         |||      ||||||||||||||
Db    788 GATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC------TTGAGACTAGTAGGA 841

Qy    267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
          |||   |||   |||   ||||||:::|||:::         :::|||   ||||||:::|||
Db    842 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGTC 901
```

```
Qy     287 CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
               |||         |||                  |||||||||||||||||||||||||||
Db     902 TGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGGG 961

Qy     307 ThrAlaLeu------HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
              :::|||         |||                     |||           :::
Db     962 AAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCATC 1021

Qy     325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
           |||||||||   |||   |||||||||   |||      |||      |||:::|||
Db    1022 TGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAGA 1081

Qy     345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
                          ||||||   ||||||   |||||:::||||||||||||
Db    1082 TTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCA 1135
```

RESULT 13
US-09-907-824-147
; Sequence 147, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 147
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-824-147

```
Alignment Scores:
Pred. No.:                 2.64e-60      Length:        1686
Score:                     691.50        Matches:       148
Percent Similarity:        52.51%        Conservative:  40
Best Local Similarity:     41.34%        Mismatches:    139
Query Match:               8.50%         Indels:        31
DB:                        9             Gaps:          9

US-09-759-130B-381 (1-1453) x US-09-907-824-147 (1-1686)

Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----------Phe 36
           |||||||||                  |||:::|||      |||           |||
Db     107 CTATTCTCC-----------------TTGATCCTTGCCATTTGCACCAGACCTGGATTC 148

Qy      37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
           |||   |||                            :::|||||||||   |||
Db     149 CTAGCGTCTCCATCTGGA--------------GTGCGGCTGGTGGGGGGCCTCCACCGC 193

Qy      57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
           |||   |||   |||||||||:::   :::|||||||||||||||||||||||||||||||
Db     194 TGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATGACGGC 253

Qy      77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
           |||:::                  |||:::|||::::::|||||||||         :::
Db     254 TGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313

Qy      97 Met---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
                       :::                  :::      |||:::   :::
Db     314 ACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCAA 373

Qy     114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
              |||||||||   |||   |||         |||   :::|||:::         |||
Db     374 TCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAA------ 427

Qy     134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------------------ 147
              ::::::|||   |||   ||||||   |||   :::|||
Db     428 GTTTATGATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCT 487

Qy     148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
           :::                  |||:::||||||   ||||||            |||   ||||||
Db     488 TTCTCCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGC 547

Qy     168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
           ||||||||||||   |||   :::|||   |||:::|||         ||||||:::|||
Db     548 GTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGG 607

Qy     188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
              |||   ||||||||||||||||||||||||         :::   ::::::
Db     608 GCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGC 667

Qy     208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
```

```
                 |||        |||       |||    |||||||||||         :::    |||    |||
Db     668 TGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727

Qy     227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
              |||        |||   :::|||          ||||||    :::    |||:::|||
Db     728 CGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGAAGAACACCTGCAACCAT 787

Qy     247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
              :::||||||      :::    |||   |||       |||       ||||||||||||||
Db     788 GATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC------TTGAGACTAGTAGGA 841

Qy     267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
           |||   |||   |||   ||||||:::|||:::       :::|||   ||||||:::|||
Db     842 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGTC 901

Qy     287 CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
           |||         |||                  ||||||||||||||||||||||||||
Db     902 TGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGGG 961

Qy     307 ThrAlaLeu------HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
              :::|||             |||                   |||           :::
Db     962 AAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCATC 1021

Qy     325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
           |||||||||   |||   |||||||||    |||       |||       |||:::|||
Db    1022 TGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAGA 1081

Qy     345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
                             ||||||    ||||||    ||||||:::||||||||||||
Db    1082 TTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCA 1135
```

RESULT 14
US-09-907-841-147
; Sequence 147, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 147
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-841-147

```
Alignment Scores:
Pred. No.:              2.64e-60     Length:        1686
Score:                  691.50       Matches:       148
Percent Similarity:     52.51%       Conservative:  40
Best Local Similarity:  41.34%       Mismatches:    139
Query Match:            8.50%        Indels:        31
DB:                     9            Gaps:          9
```

US-09-759-130B-381 (1-1453) x US-09-907-841-147 (1-1686)

```
Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys------------Phe 36
           |||||||||                   |||:::|||        |||          |||
Db     107 CTATTCTCC------------------TTGATCCTTGCCATTTGCACCAGACCTGGATTC 148

Qy      37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
           |||   |||                           :::|||||||||    |||
Db     149 CTAGCGTCTCCATCTGGA---------------GTGCGGCTGGTGGGGGGCCTCCACCGC 193

Qy      57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
           |||   |||   |||||||||:::      :::||||||||||||||||||||||||||||
Db     194 TGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATGACGGC 253

Qy      77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
           |||:::                |||:::|||::::::|||||||||        :::
Db     254 TGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313

Qy      97 Met---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
                           :::                :::        |||:::    :::
Db     314 ACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCAA 373

Qy     114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
              |||||||||    |||    |||       |||    :::|||:::     |||
Db     374 TCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAA------ 427

Qy     134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----------------- 147
              ::::::|||    |||    ||||||    |||    :::|||
Db     428 GTTTATGATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCT 487

Qy     148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
           :::                |||:::||||||    ||||||          |||   ||||||
Db     488 TTCTCCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGC 547

Qy     168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
           ||||||||||||    |||    :::|||    |||:::|||         ||||||:::|||
Db     548 GTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGG 607

Qy     188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
              |||   ||||||||||||||||||||||||          :::    ::::::
Db     608 GCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGAGGGCTGTACTGACTCAAAAACGC 667

Qy     208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
              |||       |||       |||    ||||||||||||          :::    |||    |||
```

```
Db     668 TGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727

Qy     227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
              |||      |||    :::|||           ||||||    :::   |||:::|||
Db     728 CGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCAT 787

Qy     247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
           :::||||||     :::   |||   |||       |||       ||||||||||||||
Db     788 GATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC------TTGAGACTAGTAGGA 841

Qy     267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
            |||   |||   |||   ||||||:::|||:::      :::|||   ||||||:::|||
Db     842 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGTC 901

Qy     287 CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
            |||        |||              ||||||||||||||||||||||||||||||
Db     902 TGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGGG 961

Qy     307 ThrAlaLeu------HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
              :::|||         |||                 |||          :::
Db     962 AAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCATC 1021

Qy     325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
           |||||||||   |||   |||||||||   |||         |||      |||:::|||
Db    1022 TGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAGA 1081

Qy     345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
                          ||||||   ||||||   ||||||:::||||||||||||
Db    1082 TTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCA 1135
```

RESULT 15
US-09-904-011-147
; Sequence 147, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 147
;   LENGTH: 1686
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-904-011-147

```
Alignment Scores:
Pred. No.:                2.64e-60      Length:        1686
Score:                    691.50        Matches:       148
Percent Similarity:       52.51%        Conservative:  40
Best Local Similarity:    41.34%        Mismatches:    139
Query Match:              8.50%         Indels:        31
DB:                       9             Gaps:          9

US-09-759-130B-381 (1-1453) x US-09-904-011-147 (1-1686)

Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----------Phe 36
           |||||||||                   |||:::|||     |||          |||
Db     107 CTATTCTCC-----------------TTGATCCTTGCCATTTGCACCAGACCTGGATTC 148

Qy      37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
           |||   |||                             :::|||||||||   |||
Db     149 CTAGCGTCTCCATCTGGA--------------GTGCGGCTGGTGGGGGGCCTCCACCGC 193

Qy      57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
           |||   |||   |||||||||:::   :::||||||||||||||||||||||||||||||
Db     194 TGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATGACGGC 253

Qy      77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
           |||:::                  |||:::|||:::::::|||||||||       :::
Db     254 TGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313

Qy      97 Met---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
                          :::                  :::      |||:::    :::
Db     314 ACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCAA 373

Qy     114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
              |||||||||   |||   |||         |||   :::|||:::       |||
Db     374 TCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAA------ 427

Qy     134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----------------- 147
              ::::::|||   |||   ||||||   |||   :::|||
Db     428 GTTTATGATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCT 487
```

```
Qy     148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
           :::                 |||:::||||||   ||||||          |||   ||||||
Db     488 TTCTCCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGC 547

Qy     168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
           ||||||||||||   |||   :::|||   |||:::|||      ||||||:::|||
Db     548 GTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGG 607

Qy     188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
              |||   ||||||||||||||||||||||||         :::   ::::::
Db     608 GCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGAGGGCTGTACTGACTCAAAAACGC 667

Qy     208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
              |||      |||      |||   ||||||||||||         :::   |||   |||
Db     668 TGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727

Qy     227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
              |||      |||   :::|||            ||||||   :::   |||:::|||
Db     728 CGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCAT 787

Qy     247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
           :::||||||      :::   |||   |||      |||      |||||||||||||||
Db     788 GATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC------TTGAGACTAGTAGGA 841

Qy     267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
           |||   |||   |||   ||||||:::|||:::      :::|||   ||||||:::|||
Db     842 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGTC 901

Qy     287 CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
           |||         |||                  |||||||||||||||||||||||||||
Db     902 TGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGGG 961

Qy     307 ThrAlaLeu------HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
              :::|||            |||                  |||            :::
Db     962 AAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCATC 1021

Qy     325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
           |||||||||   |||   |||||||||      |||      |||      |||:::|||
Db    1022 TGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAGA 1081

Qy     345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
                             ||||||   ||||||   ||||||:::||||||||||||
Db    1082 TTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCA 1135
```

Search completed: May 12, 2003, 17:37:41
Job time : 376 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          May 12, 2003, 06:37:56 ; Search time 446 Seconds
                                          (without alignments)
                                          7336.667 Million cell updates/sec

Title:           US-09-759-130B-381
Perfect score:   8138
Sequence:        1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09759130/runat_06052003_123553_23343/app_query.fasta_1.1607
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759130_@CGN_1_1_301_@runat_06052003_123553_23343 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_101002:*
                 1:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	8138	100.0	4359	22	AAF45124	Human TANGO 234 OR
	2	8138	100.0	4628	22	AAF45123	Human TANGO 234 cD
	3	6107.5	75.0	3670	22	AAS01217	DNA encoding human
	4	4712	57.9	2800	22	AAS44609	Human full-length
	5	3429	42.1	3716	22	AAI58649	Human polynucleoti
	6	3429	42.1	3834	22	AAH98740	Human EST-derived
c	7	3429	42.1	3834	22	AAI60435	Human polynucleoti
c	8	3429	42.1	3834	22	AAI60436	Human polynucleoti
	9	3396.5	41.7	3811	22	AAI58650	Human polynucleoti
	10	3377	41.5	3707	24	AAS94922	Human DNA sequence
	11	3082	37.9	4308	22	AAF45141	Bovine WC1 ORF. B
	12	2435	29.9	5802	19	AAV49652	Human SRCR protein
	13	1763.5	21.7	2697	24	AAS17590	DNA encoding novel
	14	1404	17.3	820	21	AAA43017	Human secreted exp
	15	1183	14.5	690	22	AAH99258	Human protein enco
	16	1142.5	14.0	4344	24	ABK63805	Rat sequence diffe
	17	1142.5	14.0	4360	18	AAT44068	Rat von Ebner's gl
	18	1081	13.3	608	20	AAV89272	EST clone CF118.
	19	1033	12.7	562	22	AAH99326	Human protein enco
	20	972.5	12.0	1659	22	AAD16354	Human sbg14862SPER
	21	921	11.3	1269	22	AAD16353	Human sbg14862SPER
	22	830	10.2	2562	20	AAX19024	Human serine prote
	23	830	10.2	3350	20	AAV72589	Human neurotrypsin
	24	785.5	9.7	29598	19	AAV49654	Human SC2 DNA. Ho
	25	727	8.9	450	22	ABA08369	Human M160 precurs
	26	722	8.9	2376	20	AAV72590	Mouse neurotrypsin
	27	722	8.9	2614	20	AAX19027	Mouse serine prote
	28	712.5	8.8	2000	19	AAV49651	Human SRCR protein
	29	691.5	8.5	1686	20	AAX52240	Protein PRO229 cDN
	30	691.5	8.5	1686	22	AAF72398	Human PRO229 cDNA.
	31	691.5	8.5	1686	24	ABL95572	Human angiogenesis
	32	691.5	8.5	1686	24	ABL88083	Human PRO229 cDNA
	33	691.5	8.5	1686	24	ABK33592	cDNA encoding huma
	34	691.5	8.5	2181	19	AAV54605	Human scavenger re
	35	690.5	8.5	1041	19	AAV49557	Human liver cell c
	36	690.5	8.5	1591	19	AAV49556	Human liver cell c
	37	684	8.4	527	22	AAS44781	Human contig polyn
	38	647.5	8.0	3767	21	AAA96737	Polynucleotide iso
	39	639.5	7.9	3532	24	AAI67789	Human lysyl oxidas
	40	638	7.8	2420	22	AAD19222	Human CG153 (or C5
	41	635.5	7.8	5059	22	AAD19230	Human CG153 (or C5
	42	634.5	7.8	2262	21	AAA47799	Human lysyl oxidas
	43	634.5	7.8	2262	22	AAI67193	Nucleotide sequenc
	44	634.5	7.8	2785	24	AAD24786	Human secreted pro
	45	634.5	7.8	2920	21	AAA47798	Human lysyl oxidas

ALIGNMENTS

```
RESULT 1
AAF45124
ID   AAF45124 standard; cDNA; 4359 BP.
XX
AC   AAF45124;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human TANGO 234 ORF.
XX
KW   Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
DR   P-PSDB; AAB66037.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 1; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence 4359 BP; 1060 A; 881 C; 1303 G; 1115 T; 0 other;

Alignment Scores:
Pred. No.:                0            Length:          4359
Score:                    8138.00      Matches:         1453
Percent Similarity:       100.00%      Conservative:    0
Best Local Similarity:    100.00%      Mismatches:      0
Query Match:              100.00%      Indels:          0
DB:                       22           Gaps:            0

US-09-759-130B-381 (1-1453) x AAF45124 (1-4359)

Qy       1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 60

Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 120

Qy      41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180

Qy      61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240

Qy      81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 300

Qy     101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 360

Qy     121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 420

Qy     141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 480

Qy     161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 540

Qy     181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600

Qy     201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 660

Qy     221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 720

Qy     241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 780

Qy     261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 840

Qy     281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 900

Qy     301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 960

Qy     321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1020

Qy     341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1080

Qy     361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1140

Qy     381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1200

Qy     401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1260

Qy     421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1320

Qy     441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1380

Qy     461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1440

Qy     481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1441 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGACTGTG 1500

Qy    501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1560

Qy    521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AAGCCTATGCATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTG 1620

Qy    541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1680

Qy    561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1740

Qy    581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1800

Qy    601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860

Qy    621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1920

Qy    641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 1980

Qy    661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2040

Qy    681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100

Qy    701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2160

Qy    721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220

Qy    741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2280

Qy    761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2340

Qy    781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2400

Qy    801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2460

Qy    821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2520

Qy    841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2580

Qy    861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2640

Qy    881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2700

Qy    901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2760

Qy    921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2820

Qy    941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2880

Qy    961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940

Qy    981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000

Qy   1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3060

Qy   1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3120

Qy   1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3180

Qy   1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3240

Qy   1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3300

Qy   1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3360

Qy   1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3420

Qy   1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480

Qy   1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3540

Qy   1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3600

Qy   1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3660
```

```
Qy    1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3661 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3720

Qy    1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3721 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3780

Qy    1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3781 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3840

Qy    1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3900

Qy    1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3901 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3960

Qy    1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3961 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4020

Qy    1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4021 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4080

Qy    1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4081 GCACTTATTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4140

Qy    1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4141 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4200

Qy    1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4201 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4260

Qy    1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4261 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320

Qy    1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
           |||||||||||||||||||||||||||||||||||||||
Db    4321 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4359

RESULT 2
AAF45123
ID   AAF45123 standard; cDNA; 4628 BP.
XX
AC   AAF45123;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human TANGO 234 cDNA.
XX
KW   Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW   central nervous system; focal brain disorder; bipolar affective disorder;
KW   global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW   senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW   Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW   neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS   Homo sapiens.
XX
PN   WO200077239-A2.
XX
PD   21-DEC-2000.
XX
PF   24-MAY-2000; 2000WO-US14858.
XX
PR   14-JUN-1999;   99US-0333159.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR   WPI; 2001-032313/04.
DR   P-PSDB; AAB66037.
XX
PT   TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT   screening assays and diagnostic assays and for the treatment of
PT   neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT   disease -
XX
PS   Claim 1; Fig 2; 359pp; English.
XX
CC   The present invention relates to TANGO or INTERCEPT proteins and coding
CC   sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC   AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC   sequences are useful for the treatment of neurological disorders such as
CC   central nervous system (CNS) disorders, CNS-related disorders, focal
CC   brain disorders, global-diffuse cerebral disorders and other
CC   neurological and cerebrovascular disorders. The CNS disorders include
CC   Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC   lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC   autonomic function disorders such as hypertension and sleep disorders,
CC   neuropsychiatric disorders, psychoactive substance use disorders,
CC   anxiety, and bipolar affective disorder.
XX
SQ   Sequence 4628 BP; 1174 A; 925 C; 1352 G; 1177 T; 0 other;

Alignment Scores:
Pred. No.:              0           Length:         4628
Score:                  8138.00     Matches:        1453
Percent Similarity:     100.00%     Conservative:   0
Best Local Similarity:  100.00%     Mismatches:     0
Query Match:            100.00%     Indels:         0
DB:                     22          Gaps:           0

US-09-759-130B-381 (1-1453) x AAF45123 (1-4628)

Qy       1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 87

Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 147

Qy      41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207

Qy      61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGAACACTACT 267

Qy      81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 GCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTTT 327

Qy     101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 GGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAAT 387

Qy     121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 GAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGGA 447

Qy     141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 GAAGATGTTGGTGTGAACTGTTATGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 507
```

```
Qy     161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 AACAACTCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGGACTATATGT 567

Qy     181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     568 GATGATGGGTGGAACTTGAATACTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 627

Qy     201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 TCTTTTATTTCTTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCCATTTGGCTGGAT 687

Qy     221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGGA 747

Qy     241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCTT 807

Qy     261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     808 GAACTAAGGCTTGTAGGTGGAACTAACCGCTGTATGGGGAGAGTAGAGCTGAAAATCCAA 867

Qy     281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCys 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     868 GGAAGGTGGGGGACCGTATGCCACCATAAGTGGAACAATGCTGCAGCTGATGTCGTATGC 927

Qy     301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCATTTGCAGTCAGGG 987

Qy     321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     988 TCTGATGTTGTATGGCTTGATGGTGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGAC 1047

Qy     341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1048 TGCAGACATTCCGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTGATC 1107

Qy     361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1108 TGCTCAGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAACAATTGTTCAGGG 1167

Qy     381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1168 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAAT 1227

Qy     401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1228 GAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTCTTTGGCAGTCGT 1287

Qy     421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1288 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGG 1347

Qy     441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGAACATGCTTCCGA 1407

Qy     461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCGGG 1467

Qy     481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1468 GCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1527

Qy     501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1528 TGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTGGGATGTGGA 1587

Qy     521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1588 AAGCCTATGCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1647

Qy     541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1648 GATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAACACAGTGGATGG 1707

Qy     561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1708 GGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767

Qy     581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1768 TGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTT 1827

Qy     601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1828 CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887

Qy     621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACA 1947

Qy     641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1948 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGG 2007

Qy     661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2008 TCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067

Qy     681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2068 ATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127

Qy     701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2128 GGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGA 2187

Qy     721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2188 ATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2247

Qy     741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2248 AGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACT 2307

Qy     761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2308 GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAT 2367

Qy     781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2368 TTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCT 2427

Qy     801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2428 GATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGT 2487

Qy     821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAATTGTGGAGAT 2547

Qy     841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2548 GCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAA 2607

Qy     861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2608 AAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCG 2667

Qy     881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727

Qy    901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2728 CGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787

Qy    921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAG 2847

Qy    941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTT 2907

Qy    961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967

Qy    981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2968 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027

Qy   1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCT 3087

Qy   1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3088 GCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT 3147

Qy   1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATC 3207

Qy   1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267

Qy   1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTG 3327

Qy   1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3328 GATGACCTGAACTGCACAGGAACGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGG 3387

Qy   1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 GGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCC 3447

Qy   1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 TTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507

Qy   1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 AACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTG 3567

Qy   1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 TGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACA 3627

Qy   1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3628 GGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGG 3687

Qy   1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3688 CAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATC 3747

Qy   1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3748 ACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAG 3807

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3808 ATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCG 3867

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCG 3927

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3928 TTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCA 3987

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3988 TTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGAT 4047

Qy   1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4048 GCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTA 4107

Qy   1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4108 GCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTC 4167

Qy   1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4168 ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGG 4227

Qy   1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4228 GGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCA 4287

Qy   1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4288 CATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4347

Qy   1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
          |||||||||||||||||||||||||||||||||||||||
Db   4348 TCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 4386

RESULT 3
AAS01217
ID   AAS01217 standard; cDNA; 3670 BP.
XX
AC   AAS01217;
XX
DT   04-JUL-2001  (first entry)
XX
DE   DNA encoding human secreted protein, POLY8.
XX
KW   Human secreted protein; therapeutic; diagnostic; human; cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..3475
FT                   /*tag= a
FT                   /product= "Human secreted protein POLY8"
XX
PN   WO200119856-A2.
XX
PD   22-MAR-2001.
XX
PF   13-SEP-2000; 2000WO-US25106.
XX
PR   13-SEP-1999;   99US-0153629.
PR   16-SEP-1999;   99US-0154520.
PR   20-SEP-1999;   99US-0154762.
PR   13-OCT-1999;   99US-0159231.
PR   12-SEP-2000; 2000US-0659634.
XX
PA   (CURA-) CURAGEN CORP.
```

```
XX
PI   Shimkets RA,  Fernandes E,  Herrmann JL,  Liu X,  Yang M,  Boldog FL;
XX
DR   WPI; 2001-244781/25.
DR   P-PSDB; AAU00396.
XX
PT   New POLYX polypeptide useful for treating or preventing a POLYX
PT   associated disorder, e.g. cancer -
XX
PS   Claim 9; Page 25-29; 152pp; English.
XX
CC   The sequence represents the coding sequence of human secreted protein,
CC   POLY8. POLYX nucleic acids, polypeptides and antibodies to POLYX can be
CC   used for treating or preventing a POLYX associated disorder in a subject,
CC   preferably a human. These can be used in the manufacture of a medicament
CC   for treating a syndrome associated with a human disease selected from a
CC   POLYX-associated disorder, where the therapeutic is a POLYX polypeptide,
CC   a POLYX nucleotide or a POLYX antibody. They may also be used to screen
CC   for a modulator of activity, or latency, or predisposition to a POLYX
CC   associated disorder, e.g. cancer.
XX
SQ   Sequence 3670 BP; 947 A; 752 C; 1073 G; 898 T; 0 other;

Alignment Scores:
Pred. No.:              0             Length:        3670
Score:                  6107.50       Matches:       1103
Percent Similarity:     95.33%        Conservative:  0
Best Local Similarity:  95.33%        Mismatches:    3
Query Match:            75.05%        Indels:        51
DB:                     22            Gaps:          1

US-09-759-130B-381 (1-1453) x AAS01217 (1-3670)

Qy    348 AsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer--------------- 362
          |||||||||||||||||||||||||||||||||||||||||||||
Db      2 AATTTGACTGTCTTCATCAAAACGATGTGTCTGTGATCTGCTCAGACCAGTTGCTATTC 61

Qy    362 ------------------------------------------------------------ 362

Db     62 AAACTCTGGAACTTCTGGAACTATTTGCCTCTGGAAGAGTAAGGAAGGGCCATGAAGAAT 121

Qy    362 ------------------------------------------------------------ 362

Db    122 GATGCAGATGTGAAAATAAGGTTGGAAAACCGTCCCAACAGTGAAAACAGATGTGTATGG 181

Qy    363 ------------------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsn 376
                            ||||||||||||||||||||||||||||||||||||||||||
Db    182 CCAACCCACAGGAAAATGGATGGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAAC 241

Qy    377 AsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGln 396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 AATTGTTCAGGAAGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAG 301

Qy    397 AsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerVal 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 AACTGGAAGAATGAACAAGCCCTTGTGGTTTGTAAGCAGCTAGGATGTCCGTTCAGCGTC 361

Qy    417 PheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIle 436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 TTTGGCAGTCGTCGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATA 421

Qy    437 SerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArg 456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 TCTTGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAAGCAAAGCGA 481

Qy    457 ThrCysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeu 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 ACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTA 541

Qy    477 ArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGlu 496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 AGGCTTGTCGGGGCTCATAGCCCCTGTTATGGGAGATTGGAGGTGAAATACCAAGGAGAG 601

Qy    497 TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln 516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 TGGGGGACTGTGTGTCATGACAGATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAA 661

Qy    517 LeuGlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGly 536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 TTGGGATGTGGAAAGCCTATGCATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGA 721

Qy    537 ProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu 556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722 CCTATTTGGCTGGATGACGTTTCTTGCATTGGAAATGAGTCAAATATCTGGGACTGTGAA 781

Qy    557 HisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer 576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    782 CACAGTGGATGGGGAAAGCATAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCA 841

Qy    577 GlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeu 596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    842 GGTGATGCAACATGGGGCCTGAGGCTGGTGGGCGGCAGCAACCGCTGCTCGGGAAGACTG 901

Qy    597 GluValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAla 616
          |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db    902 GAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACGGTAAAGCT 961

Qy    617 AlaAlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGly 636
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    962 GCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGA 1021

Qy    637 AsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGlu 656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1022 AACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAG 1081

Qy    657 SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu 676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1082 TCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAA 1141

Qy    677 AspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySer 696
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1142 GATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGC 1201

Qy    697 SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla 716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1202 AGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCT 1261

Qy    717 AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla 736
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1262 AATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCA 1321

Qy    737 IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn 756
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1322 ATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAAT 1381

Qy    757 SerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln 776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1382 TCTGGCTGCACTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAG 1441

Qy    777 ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArg 796
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1442 ACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGG 1501

Qy    797 LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp 816
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1502 CTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGG 1561

Qy    817 ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu 836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1562 CGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTA 1621

Qy    837 AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu 856
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1622 AACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTA 1681

Qy    857 ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle 876
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 ACTTGGGCCGAAAAGTTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATT 1741

Qy        877 ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 GTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGA 1801

Qy        897 TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 TATACAGATGTCCGACTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAAC 1861

Qy        917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GTGCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTT 1921

Qy        937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 CTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGA 1981

Qy        957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 GAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTG 2041

Qy        977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 GATAACTGTCAAATGACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCT 2101

Qy        997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 GTGATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGAC 2161

Qy       1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 CCATATTTGTCTGCAGTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTC 2221

Qy       1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 CGCCTAGTGGATGGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTC 2281

Qy       1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 TGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAG 2341

Qy       1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 CTGGGCTGTGGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGG 2401

Qy       1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       2402 CCCATCTGGCTGGATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCT 2461

Qy       1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 TCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCA 2521

Qy       1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       2522 GAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAAACAGGGAGCTGTGCTGGGAGATTG 2581

Qy       1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2582 GAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATA 2641

Qy       1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 GCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCT 2701

Qy       1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2702 TTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCAT 2761

Qy       1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2762 ATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAA 2821

Qy       1237 GluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSer 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2822 GAGACCTGGATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCT 2881

Qy       1257 GlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAsp 1276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2882 GGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGAC 2941

Qy       1277 LeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeu 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2942 CTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTG 3001

Qy       1297 ArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLys 1316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3002 AGGGACGCTTCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAA 3061

Qy       1317 GlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGly 1336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 GGAAATGAGTCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGA 3121

Qy       1337 HisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSer 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 CACAAGGAAGATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCC 3181

Qy       1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 TCAGGTCATTTAGCACTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTT 3241

Qy       1377 IleLeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSer 1396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 ATTCTATTTCTCACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCA 3301

Qy       1397 ThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLys 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 ACCAGAAGGAGGGGTTCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAG 3361

Qy       1417 ArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAsp 1436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 AGAGAGGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGAT 3421

Qy       1437 AlaSerAspThrSerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 GCTAGCGACACATCGCTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 3472

RESULT 4
AAS44609
ID   AAS44609 standard; DNA; 2800 BP.
XX
AC   AAS44609;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human full-length polynucleotide sequence #34.
XX
KW   Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW   mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW   cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW   nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW   angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW   genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW   cytostatic; antirheumatic; antiarthritic; vulnerary; antiinflammatory;
KW   antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW   neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW   immunostimulant; analgesic; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200164834-A2.
XX
PD   07-SEP-2001.
XX
```

```
PF   26-FEB-2001; 2001WO-US04926.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
PR   17-JUN-2000; 2000US-0597707.
PR   14-JUL-2000; 2000US-0616807.
PR   19-SEP-2000; 2000US-0664641.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Asundi V,  Zhang J,  Zhao QA,  Ren F;
PI   Xue AJ,  Yang Y,  Wehrman T,  Wang J,  Ma Y,  Wang D,  Chen R,  Xu C;
PI   Drmanac R;
XX
DR   WPI; 2001-589862/66.
DR   P-PSDB; AAU27709.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis, treatment of
PT   cancer, neurological, inflammatory disorders and for use in arrays for
PT   detection  -
XX
PS   Claim 1; SEQ ID No 34; 153pp; English.
XX
CC   Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC   contig polynucleotides encoding polypeptides of the invention. The DNA
CC   and protein sequences are useful for the treatment, diagnosis and
CC   prevention of various types of disorder in a mammalian subject such as a
CC   human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC   such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC   as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC   diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC   disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC   chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC   Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC   disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC   bowel disease. The sequences exhibit activity relating to angiogenesis,
CC   cell proliferation, cell differentiation, stem cell growth factor,
CC   activin or inhibin. Therefore, they can be used to manipulate stem cells
CC   in culture to give rise to neuroepithelial cells that can be used to
CC   augment or replace cells damaged by illness, accidental damage or genetic
CC   disorders. The sequences may also be used for regeneration of bone,
CC   cartilage, tendons and ligaments and in tissue repair and burn healing.
CC   Note: Some sequences for this patent did not form part of the printed
CC   specification, but were obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2800 BP; 697 A; 607 C; 813 G; 683 T; 0 other;

Alignment Scores:
Pred. No.:              3.57e-293      Length:         2800
Score:                  4712.00        Matches:        850
Percent Similarity:     99.77%         Conservative:   0
Best Local Similarity:  99.77%         Mismatches:     2
Query Match:            57.90%         Indels:         0
DB:                     22             Gaps:           0

US-09-759-130B-381 (1-1453) x AAS44609 (1-2800)

Qy      602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCys 621
            |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db       37 GGACGGTGGGGCACAGCGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGT 96

Qy      622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGly 641
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       97 AGCCAGCTGGACTGCCCATCTTCTATCATTGGCATGGGTCTGGGAAACGCTTCTACAGGA 156

Qy      642 TyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSer 661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      157 TATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGATGAGTCAGATCTCTGGTCA 216

Qy      662 CysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIle 681
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      217 TGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATC 276

Qy      682 CysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGly 701
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277 TGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGA 336

Qy      702 LysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMet 721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      337 AAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGTGCTAATGGCTGGGGAATG 396

Qy      722 AsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArg 741
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      397 AACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGA 456

Qy      742 GluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGly 761
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      457 GAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCTGGCTGCACTGGA 516

Qy      762 GlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeu 781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      517 GGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCATTTA 576

Qy      782 AsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAsp 801
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      577 AATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCCAGGCTGGTTGGAGCTGAT 636

Qy      802 MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp 821
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      637 ATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCAGACACATGGCGCTCTGTCTGTGAT 696

Qy      822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla 841
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      697 TCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAATTAAACTGTGGAGATGCC 756

Qy      842 IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys 861
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      757 ATATCTCTTTCTGTGGGAGATCACTTTGGAAAAGGGAATGGTCTAACTTGGGCCGAAAAG 816

Qy      862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      817 TTCCAGTGTGAAGGGAGTGAAACTCACCTTGCATTATGCCCCATTGTTCAACATCCGGAA 876

Qy      882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg 901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      877 GACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGA 936

Qy      902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      937 CTTGTGAATGGCAAATCCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACACTGG 996

Qy      922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      997 GGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTTCTATGCAGACAGCTC 1056

Qy      942 SerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArg 961
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1057 AGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATATATTGGAGAAAGAAGTGTTCGT 1116

Qy      962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1117 GTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAAATG 1176

Qy      982 ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1177 ACAGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACAGGA 1236

Qy     1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla 1021
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1237 AGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAATGTATCTGACCCATATTTGTCTGCA 1296

Qy     1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1297 GTTCCAGAGGGCAGTGCTTTGATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGATGGG 1356
```

```
Qy   1042 AspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCys 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1357 GACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCACGACGGCTTCTGGGGCACCATCTGT 1416

Qy   1062 AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 GATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTG 1476

Qy   1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp 1101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 GCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGGCTGGAT 1536

Qy   1102 AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly 1121
          |||||||||||||||||    |||||||||||||||||||||||||||||||||||||||
Db   1537 GACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGG 1596

Qy   1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 CAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTG 1656

Qy   1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 AGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAAC 1716

Qy   1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 GGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGC 1776

Qy   1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 AGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGC 1836

Qy   1202 SerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGln 1221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAG 1896

Qy   1222 CysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThr 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 TGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACA 1956

Qy   1242 CysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIle 1261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATC 2016

Qy   1262 TrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGlu 1281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 TGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAA 2076

Qy   1282 ValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPhe 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 GTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTTT 2136

Qy   1302 GlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPhe 1321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAGTCATTT 2196

Qy   1322 LeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAla 1341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 CTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAAGATGCT 2256

Qy   1342 GlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAla 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 GGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCATTTAGCA 2316

Qy   1362 LeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeuThr 1381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 CTTATTTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTTCTGTTTATTCTATTTCTCACG 2376

Qy   1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGly 1401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 TGGTGCCGAGTTCAGAAACAAAAACATCTGCCCCTCAGAGTTTCAACCAGAAGGAGGGGT 2436

Qy   1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHis 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2437 TCTCTCGAGGAGAATTTATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGGACCCACAT 2496

Qy   1422 GlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSer 1441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497 GGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAGCGACACATCG 2556

Qy   1442 LeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
          ||||||||||||||||||||||||||||||||||||
Db   2557 CTGTTGGGAGTTCTTCCTGCCTCTGAAGCCACAAAA 2592

RESULT 5
AAI58649
ID   AAI58649 standard; cDNA; 3716 BP.
XX
AC   AAI58649;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 852.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39493.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 852; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
```

```
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3716 BP; 993 A; 730 C; 1060 G; 933 T; 0 other;

Alignment Scores:
Pred. No.:              1.2e-210      Length:        3716
Score:                  3429.00       Matches:       598
Percent Similarity:     71.27%        Conservative:  161
Best Local Similarity:  56.15%        Mismatches:    294
Query Match:            42.14%        Indels:        12
DB:                     22            Gaps:          4

US-09-759-130B-381 (1-1453) x AAI58649 (1-3716)

Qy    364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db    222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy    384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db    282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy    404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db    342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy    424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db    402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy    444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db    462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy    463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db    522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581

Qy    483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db    582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641

Qy    503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db    642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy    523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db    702 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761

Qy    543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db    762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy    563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db    822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy    583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db    882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy    603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db    942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy    623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db   1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy    643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db   1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy    663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db   1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy    683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db   1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy    703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db   1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy    723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db   1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy    743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db   1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy    763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db   1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy    783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db   1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy    803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db   1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy    823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db   1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy    843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db   1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy    863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db   1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy    883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db   1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841

Qy    903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db   1842 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1901

Qy    923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db   1902 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1961

Qy    943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db   1962 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 2021

Qy    963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db   2022 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 2081

Qy    983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db   2082 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 2141

Qy   1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
```

```
Db   2142 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATT 2201

Qy   1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
          ||||||    ||||||:::    |||:::|||         :::||||||||||||:::|||
Db   2202 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 2261

Qy   1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
             ||||||||||||||||||||||||||||||:::|||    |||||||||||||||||
Db   2262 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 2321

Qy   1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
          |||   ||||||||||||||||||||||||||||||||:::::||||||||||||   |||
Db   2322 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 2381

Qy   1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
             |||||||||   |||||||||||||||||||||:::||||||||||||||||||:::
Db   2382 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441

Qy   1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
          :::   |||   |||   ||||||   :::|||||||||   |||   ||||||||||||
Db   2442 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 2501

Qy   1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
             :::||||||||||||||||||||||||||||||||||||||||||    :::||||||
Db   2502 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 2561

Qy   1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
          |||   ||||||         |||:::|||||||||||||||||||||||||||||||||
Db   2562 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGA 2621

Qy   1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
             ||||||:::||||||:::       :::::::::      |||:::|||||||||
Db   2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 2681

Qy   1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
          ||||||||||||   :::   |||   ::::::   |||   |||   |||      |||
Db   2682 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 2741

Qy   1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
                |||||||||||||::::::||||||||||||            ::::::|||||||||
Db   2742 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 2801

Qy   1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
             |||:::|||||||||||||:::|||:::::|||||||||||||||||||||||||||
Db   2802 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 2861

Qy   1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
          ::::::::::||||||::::::   |||   |||   ||||||||||||||||||||||||
Db   2862 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 2921

Qy   1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
          |||   |||||||||||||||||||||||||||||||||||||||   :::|||:::|||
Db   2922 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981

Qy   1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
          ||||||||||||||||||||||||   ||||||   |||   :::::::|||   ||||||
Db   2982 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 3041

Qy   1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
          ||||||||||||   ||||||||:::::::::::::|||||||||||||||||||   |||
Db   3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 3101

Qy   1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
          |||||||||   ||||:::   ||||||   |||:::||||||||||||||||||||||
Db   3102 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 3161

Qy   1343 ValArgCysSerGlyGlnSerLeuLys------------------SerLeuAsnAlaSer 1356
          |||   |||:::      |||::::::                       :::      |||
Db   3162 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 3221

Qy   1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
          |||         :::   |||            |||   |||:::::::||||||   :::|||
Db   3222 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 3281

Qy   1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
          :::            |||||||||      |||   |||:::|||:::                 |||
Db   3282 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 3332

Qy   1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
          ||||||:::|||               :::   :::         :::   ||||||   :::|||
Db   3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 3392

Qy   1415 LeuLysArgGluAsp 1419
          |||         :::|||
Db   3393 CTGAATGCAGATGAT 3407
```

RESULT 6
AAH98740

```
ID   AAH98740 standard; cDNA; 3834 BP.
XX
AC   AAH98740;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Human EST-derived coding sequence SEQ ID NO: 597.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition; ss.
XX
OS   Homo sapiens.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   P-PSDB; AAM24081.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 1; Page 584-585; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a cDNA
CC   of the invention.
XX
SQ   Sequence 3834 BP; 1017 A; 756 C; 1082 G; 979 T; 0 other;
```

Alignment Scores:

Pred. No.:	1.25e-210	Length:	3834
Score:	3429.00	Matches:	598
Percent Similarity:	71.27%	Conservative:	161
Best Local Similarity:	56.15%	Mismatches:	294
Query Match:	42.14%	Indels:	12
DB:	22	Gaps:	4

US-09-759-130B-381 (1-1453) x AAH98740 (1-3834)

```
Qy     364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db     222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy     384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db     282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy     404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db     342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy     424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db     402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy     444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db     462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy     463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db     522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581

Qy     483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db     582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641

Qy     503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db     642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy     523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db     702 GTCAGTTTCTCTGGTTCATCTAATTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761

Qy     543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db     762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy     563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db     822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy     583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db     882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy     603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db     942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy     623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db    1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy     643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db    1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy     663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db    1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy     683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db    1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy     703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db    1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy     723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db    1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy     743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db    1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy     763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db    1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy     783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db    1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy     803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db    1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy     823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db    1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy     843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db    1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy     863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db    1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy     883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db    1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841

Qy     903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db    1842 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1901

Qy     923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db    1902 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1961

Qy     943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db    1962 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 2021

Qy     963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db    2022 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 2081

Qy     983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db    2082 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 2141

Qy    1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
Db    2142 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATT 2201

Qy    1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db    2202 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 2261

Qy    1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db    2262 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 2321

Qy    1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db    2322 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 2381
```

```
Qy    1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
              |||||||||   |||||||||||||||||||:::|||||||||||||||||||:::
Db    2382 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441

Qy    1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
           :::   |||   |||   |||||   :::|||||||||   |||   |||||||||||
Db    2442 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 2501

Qy    1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
              :::||||||||||||||||||||||||||||||||||||||||   :::||||||
Db    2502 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 2561

Qy    1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
           |||   ||||||            |||:::||||||||||||||||||||||||||||||
Db    2562 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGA 2621

Qy    1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
              ||||||:::||||||:::   :::::::::              |||:::|||||||||
Db    2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 2681

Qy    1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
           ||||||||||||   :::   |||   ::::::   |||   |||   |||      |||
Db    2682 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 2741

Qy    1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
                 ||||||||||||:::::||||||||||||            ::::::|||||||||
Db    2742 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 2801

Qy    1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
              |||:::|||||||||:::|||::::::||||||:::|||||||||||||||||||||
Db    2802 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 2861

Qy    1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
           :::::::::||||||::::::   |||   |||   |||||||||||||||||||||||||
Db    2862 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 2921

Qy    1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
           |||   ||||||||||||||||||||||||||||||||||||||   :::|||:::|||
Db    2922 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981

Qy    1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
           ||||||||||||||||||||||||   ||||||   |||   :::::::|||   ||||||
Db    2982 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 3041

Qy    1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
           ||||||||||||   |||||||||:::::::::::|||||||||||||||||||   |||
Db    3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 3101

Qy    1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
           |||||||||   |||:::   ||||||   |||:::|||||||||||||||||||||||
Db    3102 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 3161

Qy    1343 ValArgCysSerGlyGlnSerLeuLys------------------SerLeuAsnAlaSer 1356
           |||   |||:::      |||::::::                  :::         |||
Db    3162 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 3221

Qy    1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
           |||         :::   |||            |||   |||::::::||||||   :::|||
Db    3222 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 3281

Qy    1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
           :::      ||||||||   |||   |||:::|||:::         |||
Db    3282 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 3332

Qy    1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
           ||||||:::|||            :::   :::         :::   ||||||   :::|||
Db    3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 3392

Qy    1415 LeuLysArgGluAsp 1419
           |||      :::|||
Db    3393 CTGAATGCAGATGAT 3407
```

RESULT 7

```
AAI60435/c
ID   AAI60435 standard; cDNA; 3834 BP.
XX
AC   AAI60435;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4424.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41279.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4424; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

Alignment Scores:
Pred. No.:              1.25e-210     Length:         3834
Score:                  3429.00       Matches:        598
Percent Similarity:     71.27%        Conservative:   161
Best Local Similarity:  56.15%        Mismatches:     294
Query Match:            42.14%        Indels:         12
DB:                     22            Gaps:           4
```

US-09-759-130B-381 (1-1453) x AAI60435 (1-3834)

```
Qy    364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db   3613 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 3554

Qy    384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db   3553 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 3494

Qy    404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db   3493 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 3434

Qy    424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db   3433 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 3374

Qy    444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db   3373 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 3314

Qy    463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db   3313 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 3254

Qy    483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db   3253 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 3194

Qy    503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db   3193 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 3134

Qy    523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db   3133 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 3074

Qy    543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db   3073 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 3014

Qy    563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db   3013 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 2954

Qy    583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db   2953 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 2894

Qy    603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db   2893 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 2834

Qy    623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db   2833 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 2774

Qy    643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db   2773 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 2714

Qy    663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db   2713 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 2654

Qy    683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db   2653 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 2594

Qy    703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db   2593 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 2534

Qy    723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db   2533 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 2474

Qy    743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db   2473 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 2414

Qy    763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db   2413 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 2354

Qy    783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db   2353 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 2294

Qy    803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db   2293 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 2234

Qy    823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db   2233 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 2174

Qy    843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db   2173 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 2114

Qy    863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db   2113 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 2054

Qy    883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db   2053 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1994

Qy    903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db   1993 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1934

Qy    923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db   1933 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1874

Qy    943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db   1873 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 1814

Qy    963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db   1813 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 1754

Qy    983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db   1753 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 1694

Qy   1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
Db   1693 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATT 1634

Qy   1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db   1633 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 1574

Qy   1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db   1573 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 1514

Qy   1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db   1513 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 1454
```

```
Qy   1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
             |||||||||   ||||||||||||||||||||:::|||||||||||||||||:::
Db   1453 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 1394

Qy   1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
          :::   |||   |||   |||||   :::||||||||   |||   |||||||||||
Db   1393 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 1334

Qy   1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
             :::|||||||||||||||||||||||||||||||||||||||||   :::||||||
Db   1333 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 1274

Qy   1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
          |||   ||||||         |||:::||||||||||||||||||||||||||||||||
Db   1273 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTACAATGGA 1214

Qy   1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
             ||||||:::||||||:::   :::::::::         |||:::||||||||
Db   1213 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 1154

Qy   1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
          |||||||||||    :::   |||   ::::::   |||   |||   |||      |||
Db   1153 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 1094

Qy   1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
                   |||||||||||::::::||||||||||||         ::::::||||||||
Db   1093 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 1034

Qy   1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
             |||:::||||||||:::|||::::::||||||:::||||||||||||||||||||
Db   1033 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 974

Qy   1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
          :::::::::||||||::::::   |||   |||   |||||||||||||||||||||||
Db    973 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 914

Qy   1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
          |||   ||||||||||||||||||||||||||||||||||||||   :::|||:::|||
Db    913 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 854

Qy   1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
          ||||||||||||||||||||||||   ||||||   |||   :::::::|||   ||||||
Db    853 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 794

Qy   1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
          ||||||||||||   ||||||::::::::::::|||||||||||||||||||||   |||
Db    793 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 734

Qy   1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
          |||||||||   ||||::::   ||||||   |||:::||||||||||||||||||||
Db    733 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 674

Qy   1343 ValArgCysSerGlyGlnSerLeuLys------------------SerLeuAsnAlaSer 1356
          |||   |||:::      |||::::::               :::      |||
Db    673 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 614

Qy   1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
          |||      :::   |||            |||   |||::::::||||||   :::|||
Db    613 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 554

Qy   1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
          :::         |||||||||         |||   |||:::|||:::        |||
Db    553 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 503

Qy   1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
          ||||||:::|||            :::   :::      :::   ||||||   :::|||
Db    502 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 443

Qy   1415 LeuLysArgGluAsp 1419
          |||      :::|||
Db    442 CTGAATGCAGATGAT 428
```

```
RESULT 8
AAI60436/c
ID   AAI60436 standard; cDNA; 3834 BP.
XX
AC   AAI60436;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4425.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41280.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4425; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

Alignment Scores:
Pred. No.:              1.25e-210       Length:         3834
Score:                  3429.00         Matches:        598
Percent Similarity:     71.27%          Conservative:   161
Best Local Similarity:  56.15%          Mismatches:     294
Query Match:            42.14%          Indels:         12
```

```
DB:                 22          Gaps:          4
US-09-759-130B-381 (1-1453) x AAI60436 (1-3834)

Qy    364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db   3613 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 3554

Qy    384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db   3553 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 3494

Qy    404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db   3493 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 3434

Qy    424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db   3433 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 3374

Qy    444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db   3373 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 3314

Qy    463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db   3313 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 3254

Qy    483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db   3253 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 3194

Qy    503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db   3193 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 3134

Qy    523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db   3133 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 3074

Qy    543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db   3073 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 3014

Qy    563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db   3013 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 2954

Qy    583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db   2953 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 2894

Qy    603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db   2893 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 2834

Qy    623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db   2833 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 2774

Qy    643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db   2773 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 2714

Qy    663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db   2713 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 2654

Qy    683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db   2653 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 2594

Qy    703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db   2593 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 2534

Qy    723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db   2533 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 2474

Qy    743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db   2473 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 2414

Qy    763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db   2413 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 2354

Qy    783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db   2353 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 2294

Qy    803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db   2293 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 2234

Qy    823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db   2233 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 2174

Qy    843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db   2173 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 2114

Qy    863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db   2113 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 2054

Qy    883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902
Db   2053 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1994

Qy    903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db   1993 GTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1934

Qy    923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db   1933 TCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAA 1874

Qy    943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962
Db   1873 TGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATC 1814

Qy    963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db   1813 TGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACT 1754

Qy    983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db   1753 GCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAAC 1694

Qy   1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022
Db   1693 CAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATT 1634

Qy   1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db   1633 CCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGA 1574

Qy   1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db   1573 GGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGAT 1514

Qy   1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
```

```
Db   1513 GACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCC 1454

Qy   1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
                |||||||||    |||||||||||||||||||||:::|||||||||||||||||:::
Db   1453 ATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 1394

Qy   1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
          :::    |||   |||   |||||    :::|||||||||   |||    |||||||||||
Db   1393 ATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAG 1334

Qy   1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
           :::|||||||||||||||||||||||||||||||||||||||||||||||     :::||||||
Db   1333 CAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGA 1274

Qy   1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
          |||   ||||||           |||:::|||||||||||||||||||||||||||||||
Db   1273 CTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTACAATGGA 1214

Qy   1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
             ||||||:::||||||:::    :::::::::               |||:::|||||||||
Db   1213 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGG 1154

Qy   1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
          |||||||||||    :::    |||    ::::::    |||    |||    |||        |||
Db   1153 CAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCC 1094

Qy   1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
             |||||||||||||:::::::||||||||||||                 ::::::||||||||
Db   1093 ATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGC 1034

Qy   1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
             |||:::|||||||||:::|||::::::||||||||:::||||||||||||||||||||
Db   1033 CCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGT 974

Qy   1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
          :::::::::||||||:::::::    |||    |||    ||||||||||||||||||||||
Db    973 GACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGG 914

Qy   1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1282
          |||    ||||||||||||||||||||||||||||||||||||||||    :::|||:::|||
Db    913 CATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 854

Qy   1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
          ||||||||||||||||||||||||||||||    ||||||    |||    ::::::|||    ||||||
Db    853 GTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGT 794

Qy   1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
          ||||||||||||    ||||||||||:::::::::::::||||||||||||||||||||    |||
Db    793 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTG 734

Qy   1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
          |||||||||    |||:::    ||||||    |||:::||||||||||||||||||||||
Db    733 TGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCA 674

Qy   1343 ValArgCysSerGlyGlnSerLeuLys-----------------SerLeuAsnAlaSer 1356
          |||    |||:::          |||::::::                   :::          |||
Db    673 GTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCA 614

Qy   1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPhe 1376
          |||          :::    |||          |||    |||::::::||||||    :::|||
Db    613 TCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTC 554

Qy   1377 Ile------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
          :::             |||||||||          |||    |||:::|||:::          |||
Db    553 GTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CTTGCA 503

Qy   1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
          ||||||:::|||                :::    :::       :::    ||||||    :::|||
Db    502 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGC 443

Qy   1415 LeuLysArgGluAsp 1419
          |||       :::|||
Db    442 CTGAATGCAGATGAT 428
```

```
RESULT 9
AAI58650
ID   AAI58650 standard; cDNA; 3811 BP.
XX
AC   AAI58650;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 853.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39494.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 853; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3811 BP; 1018 A; 751 C; 1081 G; 961 T; 0 other;

Alignment Scores:
Pred. No.:              1.51e-208     Length:         3811
Score:                  3396.50       Matches:        597
Percent Similarity:     69.03%        Conservative:   161
Best Local Similarity:  54.37%        Mismatches:     295
```

```
Query Match:            41.74%       Indels:       45
DB:                     22           Gaps:         5

US-09-759-130B-381 (1-1453) x AAI58650 (1-3811)

Qy    364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db    222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGAGAGTGGAA 281

Qy    384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db    282 GTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTC 341

Qy    404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db    342 TCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAAT 401

Qy    424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db    402 TCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTCA 461

Qy    444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db    462 GCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACAA 521

Qy    463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db    522 GATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGG 581

Qy    483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db    582 AATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGAT 641

Qy    503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db    642 GATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGCT 701

Qy    523 MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db    702 GTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGAT 761

Qy    543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db    762 CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAAG 821

Qy    563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db    822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAGC 881

Qy    583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db    882 CTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGA 941

Qy    603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer 622
Db    942 GAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAG 1001

Qy    623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db   1002 CAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATTT 1061

Qy    643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db   1062 GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATGT 1121

Qy    663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db   1122 AAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181

Qy    683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db   1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACA 1241

Qy    703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db   1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAA 1301

Qy    723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db   1302 GAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAA 1361

Qy    743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db   1362 GTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAAT 1421

Qy    763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db   1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTAT 1481

Qy    783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db   1482 GAAGAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATT 1541

Qy    803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db   1542 CCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCG 1601

Qy    823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db   1602 GACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTC 1661

Qy    843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db   1662 TCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAATTC 1721

Qy    863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db   1722 CAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAAGGA 1781

Qy    883 ThrCysIleHisSerArgGluValGlyValValCysSerArg------------------ 896
Db   1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGTAAGACCCAGAAAACATCT 1841

Qy    896 ------------------------------------------------------------ 896
Db   1842 TTAATTGGTTCTTATACTGTGAAAGGGACAGGGTTAGGGAGTCATAGCTGTCTTTTTCTA 1901

Qy    897 ---------------------TyrThrAspValArgLeuValAsnGlyLysSerGlnCys 909
Db   1902 AAGCCCTGTCTCCTTCCAGGATACACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGT 1961

Qy    910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929
Db   1962 GAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGATCCCTCTGTAACTCTCACTGG 2021

Qy    930 AspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949
Db   2022 GACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACC 2081

Qy    950 ThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCys 969
Db   2082 CCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAGATCTGGAGGCATATGTTTCACTGC 2141

Qy    970 LeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCys 989
Db   2142 ACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTAACTGCTCTAGGTGCTTCATTATGT 2201

Qy    990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009
Db   2202 CCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAACCAGTCCCAAACACTGTCCTCG 2261

Qy   1010 CysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIle 1029
Db   2262 TGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCC 2321

Qy   1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049
```

```
             |||:::|||        :::|||||||||||:::|||       ||||||||||||||||
Db     2322 TGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTA 2381

Qy     1050 GluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069
             ||||||||||||:::|||    |||||||||||||||||||    ||||||||||||||
Db     2382 GAGATCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAGTGAT 2441

Qy     1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089
             |||||||||||||||:::::|||||||||||    |||  :|||||||||   ||||||
Db     2442 GCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGGTTCTGCT 2501

Qy     1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109
             ||||||||||||||||||||||||||||||||||||::::::   |||    |||    |||
Db     2502 CATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAAATGCAATGGAAAAGAA 2561

Qy     1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129
             |||    :::|||||||||    |||    ||||||||||||    :::|||||||||||||||
Db     2562 TCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAATTGCAGGCACAAGGAG 2621

Qy     1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149
             |||||||||||||||||||||||||||    :::|||||||||    ||||||
Db     2622 GATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTGAGACTGACCAGTGAAGCCAGCAGA 2681

Qy     1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169
             |||:::|||||||||||||||||||||||||||||||||    ||||||:::||||||:::
Db     2682 GAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTGGGGCACTGTTGGCAAG 2741

Qy     1170 ArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
                :::::::::              |||:::|||||||||||||||||||||    :::
Db     2742 AGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCAGACAAA 2801

Qy     1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
             |||   ::::::   |||   |||   |||        |||      ||||||||||||:::
Db     2802 GGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTGGACAAT 2861

Qy     1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
             :::||||||||||||        ::::::|||||||||   |||:::|||||||||:::
Db     2862 GTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGGGAGAAG 2921

Qy     1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
             |||::::::||||||||||||||||||||||||::::::::::::::::::||||||:::::
Db     2922 AGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAACAAGATAAGACTTCAG 2981

Qy     1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
                |||   |||   ||||||||||||||||||||||||||   ||||||||||||||
Db     2982 GAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGAGGTTCCTGGGGACA 3041

Qy     1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
             ||||||||||||||||||||||||    :::|||:::||||||||||||||||||||||||
Db     3042 GTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTTGGCTGT 3101

Qy     1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
             |||   ||||||   |||   ::::::|||   |||||||||||||||||   ||||||
Db     3102 GGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCGATATGG 3161

Qy     1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
             |||:::::::::::::|||||||||||||||||   ||||||||||||   |||:::
Db     3162 CTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGATTGTCCTGCCAGACGC 3221

Qy     1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
             ||||||   |||:::||||||||||||||||||||   |||   |||:::      |||
Db     3222 TGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAATTGCACAGATATTTCA 3281

Qy     1350 LeuLys-----------------SerLeuAsnAlaSerSerGlyHisLeuAlaLeuIle 1363
             ::::::                 :::         ||||||        :::   |||
Db     3282 GTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCCCGTCAGTCATCCTTTATT 3341

Qy     1364 LeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIle------LeuPheLeuThr 1381
                   |||   |||::::::||||||   :::|||:::            ||||||||
Db     3342 GCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTCGTCGCATTATTCTTCTTGACT 3401

Qy     1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGly 1401
                   |||   |||:::|||:::         |||   ||||||:::|||
Db     3402 AAAAAGCGAAGACAGAGACAGCGG---------CTTGCAGTTTCCTCAAGAGGAGAGAAC 3452

Qy     1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAsp 1419
                :::   :::      :::   ||||||   :::||||||      :::|||
Db     3453 TTAGTCCACCAAATTCAATACCGGGAGATGAATTCTTGCCTGAATGCAGATGAT 3506
```

RESULT 10
AAS94922
ID   AAS94922 standard; DNA; 3707 BP.
XX
AC   AAS94922;
XX
DT   14-FEB-2002 (first entry)
XX
DE   Human DNA sequence #177 expressed during foam cell differentiation.
XX
KW   Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW   cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177389-A2.
XX
PD   18-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US11128.
XX
PR   05-APR-2000; 2000US-195106P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI   Tai J;
XX
DR   WPI; 2002-010925/01.
XX
PT   Composition useful for diagnosis of conditions, disorders or diseases
PT   associated with atherosclerosis, comprises several polynucleotides that
PT   are differentially expressed in foam cell development -
XX
PS   Claim 1; Page 232-233; 315pp; English.
XX
CC   The present invention relates to the isolation of human polynucleotide
CC   sequences that are differentially expressed during foam cell
CC   differentiation. The polynucleotide sequences of the invention or a
CC   composition comprising these polynucleotides are useful as a high
CC   throughput method for detecting altered expression of one or more
CC   polynucleotides in a sample. The polynucleotides can be used in the
CC   diagnosis of disorders associated with foam cell development such as
CC   atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC   coronary artery disease. The polynucleotide sequences can also be used
CC   as PCR primers and probes. The polynucleotides of the invention are also
CC   useful in gene therapy. AAS94746-AAS95021 represent the human
CC   polynucleotide sequences of the invention which are differentially
CC   expressed during foam cell differentiation.
XX
SQ   Sequence 3707 BP; 984 A; 731 C; 1062 G; 930 T; 0 other;

```
Alignment Scores:
Pred. No.:              2.61e-207      Length:        3707
Score:                  3377.00        Matches:       598
Percent Similarity:     71.00%         Conservative:  161
Best Local Similarity:  55.94%         Mismatches:    294
Query Match:            41.50%         Indels:        16
DB:                     24             Gaps:          4

US-09-759-130B-381 (1-1453) x AAS94922 (1-3707)

Qy      364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSer-GlyArgValGl 383
               |||   |||   ||||||||||||   ||||||   |||   ||||||  ||||||||||
Db      222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGGGAGAGTGGA 281

Qy      383 uValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAl 403
```

```
Db    282 AGTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGT 341

Qy    403 aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLy 423
Db    342 CTCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGATGGGCTAA 401

Qy    423 sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe 443
Db    402 TTCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGGAATGAGTC 461

Qy    443 rAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSe 462
Db    462 AGCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACTCACCAACA 521

Qy    462 rAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHi 482
Db    522 AGATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGG 581

Qy    482 sSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHi 502
Db    582 GAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACAGTGTGTGA 641

Qy    502 sAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPr 522
Db    642 TGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGC 701

Qy    522 oMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAs 542
Db    702 TGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGA 761

Qy    542 pValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLy 562
Db    762 TCTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGATGGGGAAA 821

Qy    562 sHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGl 582
Db    822 GCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTGAG 881

Qy    582 yLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGl 602
Db    882 CCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGG 941

Qy    602 yArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSe 622
Db    942 AGAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAA 1001

Qy    622 rGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTy 642
Db   1002 GCAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATT 1061

Qy    642 rGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCy 662
Db   1062 TGGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTCTGGCAATG 1121

Qy    662 sArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCy 682
Db   1122 TAAACACCATGAATGGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATG 1181

Qy    682 sSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLy 702
Db   1182 TTCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGAC 1241

Qy    702 sValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAs 722
Db   1242 AGTTGAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAA 1301

Qy    722 nIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGl 742
Db   1302 AGAAGCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCA 1361

Qy    742 uProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGl 762
Db   1362 AGTGTACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAA 1421

Qy    762 yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs 782
Db   1422 TGAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTA 1481

Qy    782 nMetGluAlaSerLeuIleCysSer-AlaHisArg-GlnProArgLeuValGlyAlaAsp 801
Db   1482 TGAAGAAGCCAAAATTACCTGCTCAGGCCCACAGGGGAACCCAGACTGGTTGGAGGGGAC 1541

Qy    802 MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp 821
Db   1542 ATTCCCTGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGAT 1601

Qy    822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla 841
Db   1602 TCAGACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTT 1661

Qy    842 IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys 861
Db   1662 GTCTCTATCCTGGGGGGAGCTCACTTTGGAGAGGGAAATGGACAGATCTGGGCTGAAGAA 1721

Qy    862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881
Db   1722 TTCCAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAGTAGCACCCCGCCCAGAA 1781

Qy    882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg 901
Db   1782 GGAACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGC 1841

Qy    902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921
Db   1842 TTGGTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGGTGCCTGG 1901

Qy    922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941
Db   1902 GGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTT 1961

Qy    942 SerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArg 961
Db   1962 AAATGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAAGGAAATGGTCAG 2021

Qy    962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981
Db   2022 ATCTGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGTCCTGTA 2081

Qy    982 ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001
Db   2082 ACTGCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGA 2141

Qy   1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla 1021
Db   2142 AACCAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACC 2201

Qy   1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041
Db   2202 ATTCCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGA 2261

Qy   1042 AspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCys 1061
Db   2262 GGAGGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCATCTGT 2321

Qy   1062 AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081
Db   2322 GATGACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAG 2381

Qy   1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp 1101
Db   2382 GCCATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGAT 2441

Qy   1102 AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly 1121
Db   2442 GAGATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGG 2501
```

```
Qy    1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141
            |||   ::::||||||||||||||||||||||||||||||||||||||||||    :::|||
Db    2502 CAGCAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGTCTCTG 2561

Qy    1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
            ||||.||   ||||||           |||:::||||||||||||||||||||||||||||
Db    2562 AGACTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAAT 2621

Qy    1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181
            |||    ||||||:::||||||:::    :::::::::            |||:::||||||
Db    2622 GGAGCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGC 2681

Qy    1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201
            ||||||||||||||    :::    |||    ::::::    |||    |||    |||
Db    2682 AGGCAGCTGGGCTGTGCAGACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATG 2741

Qy    1202 SerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp-Gl 1221
            |||       ||||||||||||::::::||||||||||||            :::::::||| ||
Db    2742 TCCATTCCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGGCA 2801

Qy    1221 nCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleTh 1241
            ||||    |||:::||||||||||:::|||::::::||||||:::||||||||||||||||
Db    2802 GTGCCCATCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCAC 2861

Qy    1241 rCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIl 1261
            ||||:::::::::||||||::::::    |||    |||    ||||||||||||||||||
Db    2862 ATGTGACAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGAT 2921

Qy    1261 eTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGl 1281
            ||||||    ||||||||||||||||||||||||||||||||||||||    ::::|||::
Db    2922 CTGGCATGGAGGTTCCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCA 2981

Qy    1281 uValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPh 1301
            :|||||||||||||||||||||||||||    ||||||    |||    :::::::|||    ||
Db    2982 GGTGGTGTGTCAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTT 3041

Qy    1301 eGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPh 1321
            |||||||||||||||    ||||||||::::::::::::||||||||||||||||
Db    3042 TGGTCAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTC 3101

Qy    1321 eLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAl 1341
            ||||||||||||    |||:::    ||||||    |||:::|||||||||||||||||||||
Db    3102 CTTGTGGGATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGC 3161

Qy    1341 aGlyValArgCysSerGlyGlnSerLeuLys-----------------SerLeuAsnAl 1355
            |    |||    |||:::    |||:::::                    :::
Db    3162 TGCAGTGAATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCG 3221

Qy    1355 aSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLe 1375
            ||||||          :::    |||          |||    |||::::::||||||    ::
Db    3222 CTCATCCCGTCAGTCATCCTTTATTGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCAT 3281

Qy    1375 uPheIle------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLe 1393
            :|||:::          |||||||||    |||    |||:::|||:::          ||
Db    3282 TTTCGTCGCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCGG---------CT 3332

Qy    1393 uArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluTh 1413
            |    ||||||:::|||                 :::    :::       :::    ||||||    ::
Db    3333 TGCAGTTTCCTCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAATTC 3392

Qy    1413 rCysLeuLysArgGluAsp 1419
            :||||||        :::|||
Db    3393 TTGCCTGAATGCAGATGAT 3411

RESULT 11
AAF45141
ID    AAF45141 standard; cDNA; 4308 BP.
XX
AC    AAF45141;
XX
DT    30-MAR-2001  (first entry)
XX
DE    Bovine WC1 ORF.
XX
KW    Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW    central nervous system; focal brain disorder; bipolar affective disorder;
KW    global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW    senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW    Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW    neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS    Bos sp.
XX
PN    WO200077239-A2.
XX
PD    21-DEC-2000.
XX
PF    24-MAY-2000; 2000WO-US14858.
XX
PR    14-JUN-1999;   99US-0333159.
XX
PA    (MILL-) MILLENNIUM PHARM INC.
XX
PI    McCarthy SA,  Fraser CC,  Sharp JD,  Barnes TM;
XX
DR    WPI; 2001-032313/04.
DR    P-PSDB; AAB66088.
XX
PT    TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT    screening assays and diagnostic assays and for the treatment of
PT    neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT    disease -
XX
PS    Disclosure; Fig 2; 359pp; English.
XX
CC    The present invention relates to TANGO or INTERCEPT proteins and coding
CC    sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC    AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC    sequences are useful for the treatment of neurological disorders such as
CC    central nervous system (CNS) disorders, CNS-related disorders, focal
CC    brain disorders, global-diffuse cerebral disorders and other
CC    neurological and cerebrovascular disorders. The CNS disorders include
CC    Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC    lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC    autonomic function disorders such as hypertension and sleep disorders,
CC    neuropsychiatric disorders, psychoactive substance use disorders,
CC    anxiety, and bipolar affective disorder. The present sequence is a
CC    sequence used in a sequence homology alignment with the TANGO/INTERCEPT
CC    sequences of the present invention.
XX
SQ    Sequence 4308 BP; 866 A; 1093 C; 1410 G; 939 T; 0 other;

Alignment Scores:
Pred. No.:               2.74e-188       Length:          4308
Score:                   3082.00         Matches:         613
Percent Similarity:      55.85%          Conservative:    208
Best Local Similarity:   41.70%          Mismatches:      495
Query Match:             37.87%          Indels:          154
DB:                      22              Gaps:            23

US-09-759-130B-381 (1-1453) x AAF45141 (1-4308)

Qy      28 CysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu 47
            |||:::|||||||||                             :::        |||    ||||||
Db      37 TGTGTCCTCCTCCTCGGC---------------ACCATGGTGGGTGGTCAAGCTCTGGAG 81

Qy      48 LeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
            |||||||||    :::|||          |||    |||    ||||||||||||    ||||||
Db      82 CTGAGGTTGAAGGATGGAGTCCATCGCTGTGAGGGGAGAGTGGAAGTGAAGCACCAAGGA 141

Qy      68 GlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrValValCysLys 87
            :::||||||||||||                |||                :::::::||||||||:::
Db     142 GAATGGGGCACAGTGGATGGTTACAGGTGGACATTGAAGGATGCATCTGTAGTGTGCAGA 201

Qy      88 GlnLeuGlyCysProPheSerPheAlaMet---------PheArgPheGlyGlnAlaVal 104
```

```
               |||||||||||     :::                  ||||||      :::
Db      202 CAGCTGGGGTGTGGAGCTGCCATTGGTTTTCCTGGAGGGGCTTATTTGGCCAGGACTT 261

Qy      105 ThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeu 124
                     |||  ||||||||         ||||||  |||  ||||||  :::
Db      262 ---------GGCCCCATTTGGCTTTTGTATACTTCATGTGAAGGGACAGAGTCAACTGTC 312

Qy      125 TrpGluCysGlnHis---------ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141
              ::|||::|||            :::::::::  :::           ||||||
Db      313 AGTGACTGTGAGCATTCTAATATTAAAGACTATCGTAATGATGGCTATAATCATGGTCGG 372

Qy      142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161
            |||  ||||||  |||  |||              :::|||||||     |||:::
Db      373 GATGCTGGAGTAGTCTGCTCAGGA------------TTTGTGCGTCTGGCTGGAGGGGAT 420

Qy      162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
                  ||||||||||||||||||            |||  |||      :::  |||
Db      421 GGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCAGTGTCTGAT 480

Qy      182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
               :::  |||  ||||||  :::::|||  :::||||||||          :::
Db      481 GGGAACTTCACACTTGCCACTGCCCAGATCATCTGTGCAGAGTTGGGTTGTGGCAAGGCT 540

Qy      202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
                        |||                      :::|||          ::::::
Db      541 GTGTCTGTCCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGTCTGGGCTGAAGAG 600

Qy      222 IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241
                  |||:::|||  |||     ||||||  |||
Db      601 TTCAGGTGTGAGGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCCTGTCCAGGG 660

Qy      242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu 261
                  |||  |||:::            :::  |||        |||::::::
Db      661 GGCACGTGTCACCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGAAGTCCGG 720

Qy      262 LeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281
            |||   :::       |||:::::::::|||  |||:::|||||:::   |||   |||
Db      721 CTC---ATGACAAACGGCTCCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTGGA 777

Qy      282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLys 301
            :::|||      :::|||         |||:::  |||   |||:::|||:::|||:::
Db      778 CAATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTTATCTGTCGT 837

Qy      302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321
            ||||||||||||||   |||:::         |||   |||||||||
Db      838 CAGCTCGGCTGTGGAGTTGCCATCTCCACCCCCGGAGGACCACACTTGGTGGAAGAAGGT 897

Qy      322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
            |||   :::                  |||||||||  ||||||||||||||   |||
Db      898 GATCAGATCCTAACAGCCCGATTTCACTGCTCTGGGGCTGAGTCCTTCCTGTGGAGTTGT 957

Qy      342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
               :::                 ||||||  |||  |||      ||||||||||||
Db      958 CCTGTGACTGCCCTGGGTGGTCCTGACTGTTCCCATGGCAACACAGCCTCTGTGATCTGC 1017

Qy      362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
            |||
Db     1018 TCAGGAAAC--------------------------------------------------- 1026

Qy      382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
               ::::::::::::     |||
Db     1027 ---CAGATCCAGGTGCTTCCCCAG------------------------------------ 1047

Qy      402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
                                          |||   ||||||
Db     1048 ------------------------------TGCAACGACTCCGTG--------------- 1062

Qy      422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
            ::::::|||:::     |||
Db     1063 TCTCAACCTACAGGCTCTGCGGCCTCA--------------------------------- 1089

Qy      442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461

Db     1089 ------------------------------------------------------------ 1089

Qy      462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
               |||:::         |||||||||               |||||||||||
Db     1090 GAGGACAGCGCCCCCTACTGCTCAGACAGCAGG------CAGCTCCGCCTGGTGGACGGG 1143

Qy      482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
                  ||||||   ||||||:::|||:::         ||||||  |||||||||::::|||
Db     1144 GGCGGTCCCTGCGCCGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGT 1203

Qy      502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
               |||   |||·         :::|||   |||||||||:::|||||||||||||:::
Db     1204 GATGACGGCTGGGACCTGGACGATGCCCGCGTGGTGTGCAGGCAGCTGGGCTGTGGAGAA 1263

Qy      522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
               :::::      |||      :::|||      |||||||||||||||||||
Db     1264 GCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGAC 1323

Qy      542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
            :::::::::|||   |||   ||||||:::::::|||   |||         ||||||||
Db     1324 AACTTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCTTCCCGGGGCTGGGGG 1383

Qy      562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
            :::|||||||||   |||::::::|||         |||   ||||||
Db     1384 CAGCACAACTGCAGACACAAGCAGGACGCGGGGGTCATCTGCTCA------GAGTTCCTG 1437

Qy      582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601
               ||||||:::|||            :::|||:::|||   ||||||||::::::
Db     1438 GCCCTCAGGATGGTGAGTGAGGACCAGCAGTGTGCTGGGTGGCTGGAAGTTTTCTACAAT 1497

Qy      602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCys 621
            |||   ||||||:::||||||                     :::      ::::|||
Db     1498 GGGACCTGGGGCAGTGTCTGCCGTAACCCCATGGAAGACATCACTGTGTCCACGATCTGC 1557

Qy      622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
               ||||||   |||      |||            |||      |||:::|||
Db     1558 AGACAGCTTGGCTGTGGGGACAGT------------GGAACCCTCAACTCTTCTGTTGCT 1605

Qy      641 ---------GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
                     |||:::         |||:::|||   :::   |||         ::::::
Db     1606 CTTAGAGAAGGTTTTAGGCCACAGTGGGTGGATAGAATCCAGTGTCGGAAAACTGACACC 1665

Qy      658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
            ||||||   |||   :::      |||      ||||||         |||:::
Db     1666 TCTCTCTGGCAGTGTCCTTCTGACCCTTGGAATTACAACTCATGCTCTCCAAAGGAGGAA 1725

Qy      678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697
               :::      |||:::|||:::         :::::||||||||   |||
Db     1726 GCCTATATCTGGTGTGCAGACAGC------AGACAGATCCGCCTGGTGGATGGAGGTGGT 1779

Qy      698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
            ||||||:::|||:::|||||||:::         ||||||:::   |||   :::|||   :::
Db     1780 CGCTGCTCTGGGAGAGTGGAGATCCTTGACCAGGGCTCCTGGGGCACCATCTGTGATGAC 1839

Qy      718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIle 737
               |||   ::::::   |||   |||||||||:::||||||   ||||||   |||:::
Db     1840 CGCTGGGACCTGGACGATGCCCGTGTGGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTG 1899

Qy      738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
               :::               |||         :::   |||   :::
Db     1900 GACGCCACTGTCTCTTCCTTCTTCGGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTG 1959

Qy      758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
               |||   |||   |||:::   :::|||   |||      |||   |||:::|||
Db     1960 AACTGCAGAGGAGAGGAGTCCCAAGTATGGAGGTGCCCTTCCTGGGGATGGCGGCAACAC 2019

Qy      778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
               |||:::         :::|||   :::|||||||||                |||||
Db     2020 AACTGCAATCATCAAGAAGATGCAGGAGTCATCTGCTCAGGATTT------GTGCGTCTG 2073

Qy      798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
               |||   |||   ||||||||||||||||||||||||            :::   |||
Db     2074 GCTGGAGGAGATGGACCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCTGGACC 2133
```

```
Qy    818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
              |||    |||    :::|||:::|||       |||   |||:::|||   |||||
Db   2134 CCAGTGTCTGATGGAAACTTCACACTCCCCACTGCCCAGGTCATCTGTGCAGAGCTGGGA 2193

Qy    838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
          |||||    |||:::|||:::                  |||   :::   :::|||
Db   2194 TGTGGCAAGGCTGTGTCTGTCCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTC 2253

Qy    858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
          |||||||||:::|||:::|||:::|||   |||      |||      ||||||   |||
Db   2254 TGGGCTGAAGAGTTCAGGTGTGATGGGGGGGAGCCTGAGCTCTGGTCCTGCCCCAGAGTG 2313

Qy    878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
                |||      |||||||:::||||||         ||||||||||||   |||
Db   2314 CCCTGTCCAGGAGGCACATGTCTCCACAGTGGAGCTGCTCAGGTTGTCTGTTCAGTGTAC 2373

Qy    898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
          |||:::|||:::|||:::   ||||||   |||||||||:::|||||||||||||:::
Db   2374 ACAGAAGTCCAGCTTATGAAAAACGGCACCTCTCAATGTGAGGGGCAGGTGGAGATGAAG 2433

Qy    917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
          :::   |||   |||   :::||||||   :::||||||         :::|||   |||
Db   2434 ATCTCTGGACGATGGAGAGCGCTCTGTGCCTCCCACTGGAGTCTGGCCAATGCCAATGTT 2493

Qy    937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
          :::||||||||||||   ||||||   |||:::||||||      |||   ::::::
Db   2494 GTCTGTCGTCAGCTCGGCTGTGGAGTCGCCATCTCCACCCCCAGAGGACCACACTTGGTG 2553

Qy    957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
          |||         ::::::         :::||||||   |||   ||||||   |||
Db   2554 GAAGGAGGTGATCAGATCTCAACAGCCCAATTTCACTGCTCAGGGGCTGAGTCCTTCCTG 2613

Qy    977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
                :::|||   :::|||   ||||||   |||   |||   ||||||||||||   |||
Db   2614 TGGAGTTGTCCTGTGACTGCCTTGGGTGGGCCTGACTGTTCCCATGGCAACACAGCCTCT 2673

Qy    997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
          |||||||||:::|||:::   ||||||   |||      |||         :::|||
Db   2674 GTGATCTGCTCAGGAAACCACACCCAGGTGCTGCCCCAGTGCAACGACTTCCTGTCTCAA 2733

Qy   1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
          |||      ||||||      |||   |||:::      |||   :::   ::::::|||
Db   2734 CCTGCAGGCTCTGCGGCCTCAGAGGAGAGTTCTCCCTACTGCTCAGACAGCAGGCAGCTC 2793

Qy   1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
          |||||||||||||||         |||      |||||||||||||||      |||
Db   2794 CGCCTGGTGGACGGGGGCGGTCCCTGCGGCGGGAGAGTGGAGATCCTTGACCAGGGCTCC 2853

Qy   1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
          |||||||||||||||||||||      |||||||||   ||||||   |||||||||::::::
Db   2854 TGGGGCACCATCTGTGATGATGACTGGGACCTGGACGATGCCCGTGTGGTGTGCAGGCAG 2913

Qy   1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
          ||||||||||||   |||   |||||||||   |||||||||||||||   |||||||||
Db   2914 CTGGGCTGTGGAGAAGCCCTCAATGCCACGGGGTCTGCTCACTTCGGGGCAGGATCAGGG 2973

Qy   1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
          ||||||||||||||||||||||||||||||||   |||||||||:::|||:::||||||
Db   2974 CCCATCTGGCTGGACGACCTGAACTGCACAGGAAAGGAGTCCCACGTGTGGAGGTGCCCT 3033

Qy   1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
          ||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db   3034 TCCCGGGGCTGGGGGCGGCACGACTGCAGACACAAGGAGGACGCCGGGGTCATCTGCTCA 3093

Qy   1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluSerCysAlaGlyArgLeu 1156
          ||||||   |||||||||:::   |||      |||   :::   |||||||||   |||
Db   3094 GAGTTCCTGGCCCTCAGGATGGTGAGC------GAGGACCAGCAGTGTGCTGGGTGGCTG 3147

Qy   1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
          |||||||||||||||||||||||||||||||   |||      :::
Db   3148 GAGGTTTTCTACAACGGGACCTGGGGCAGTGTCTGCCGCAGCCCCATGGAAGATATCACT 3207

Qy   1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
             ::::::|||||||||||||||||||||||||||::::::|||   ::::::   :::
Db   3208 GTGTCCGTGATCTGCAGACAGCTTGGATGTGGGGACAGTGGAAGTCTCAACACCTCTGTT 3267

Qy   1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
             :::   ||||||         |||||||||   |||||||||   |||
Db   3268 GGTCTCAGGGAAGGTTCTAGACCCCGGTGGGTAGATTTAATTCAGTGTCGGAAAATGGAT 3327

Qy   1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
             |||:::|||||||||   |||   ||||||:::            ||||||   |||
Db   3328 ACCTCTCTCTGGCAGTGTCCTTCTGGCCCATGGAAATACAGTTCATGCTCTCCAAAGGAG 3387

Qy   1237 GluThrTrpIleThrCysGlu------------------------------------- 1243
          |||   :::|||:::||||||
Db   3388 GAAGCCTACATCTCATGTGAAGGAAGAAGACCCAAGAGCTGTCCAACTGCTGCCGCCTGC 3447

Qy   1244 ---------AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
                   ::::::::|||:::||||||||||||:::||||||||||||||||||
Db   3448 ACAGACAGAGAGAAGCTCCGCCTCAGGGGAGGAGACAGCGAGTGCTCAGGGCGGGTGGAG 3507

Qy   1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
          :::||||||   ||||||||||||||||||||||||||||||   ||||||||||||
Db   3508 GTGTGGCACAACGGCTCCTGGGGCACCGTGTGCGATGACTCCTGGAGCCTGGCAGAGGCT 3567

Qy   1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
          ||||||||||||||||||||||||||||||   ||||||   |||:::|||   |||:::
Db   3568 GAGGTGGTGTGTCAGCAGCTGGGCTGTGGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCA 3627

Qy   1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
          ||||||   |||      |||:::|||||||||||||:::::::::|||   |||   ||||||
Db   3628 TTTGGCCCTGGAAATGGGAGCATCTGGCTGGACGAGGTGCAGTGCGGGGGCCGGGAGTCC 3687

Qy   1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
          ||||||||||||   |||:::||||||||||||||||||||||||   |||:::||||||
Db   3688 TCCCTGTGGGACTGTGTTGCGGAGCCCTGGGGGCAGAGCGACTGCAAGCACGAGGAGGAT 3747

Qy   1341 AlaGlyValArgCysSerGly------------------------------------- 1347
          |||||||||||||||||||||
Db   3748 GCTGGTGTGAGGTGCTCTGGTGTAAGGACAACATTGCCCACGACCACAGCAGGGACCAGA 3807

Qy   1348 ---------GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeu 1364
                   ||||||      :::   :::      |||   |||   |||||||||
Db   3808 ACAACCTCAAATTCTCTCCCTGGCATCTTCTCCCTGCCTGGGGTTCTCTGCCTTATCCTG 3867

Qy   1365 SerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg 1384
             |||:::      |||:::|||:::::::|||            |||   |||   |||
Db   3868 GGGTCGCTTCTCTTCCTGGTCCTCGTCATCCTGGTGACTCAGCTACTCAGATGG---AGA 3924

Qy   1385 ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu 1404
             ::::::                  |||      :::            :::|||
Db   3925 GCAGAGCGC------------------AGAGCCTTATCCAGCTATGAAGATGCTCTTGCT 3966

Qy   1405 GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg 1424
          |||   ::::::   |||:::::::      |||   :::::::::      |||:::
Db   3967 GAAGCTGTGTATGAGGAGCTCGATTACCTTCTGACACAGAAGGAAGGTCTGGGCAGCCCA 4026

Qy   1425 ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu 1443
                   |||   |||:::         :::||||||   :::         :::
Db   4027 GATCAGATGACTGATGTCCCTGATGAAAATTATGATGATGCTGAAGAAGTACCAGTGCCT 4086

Qy   1444 GlyValLeuProAlaSerGluAlaThrLys 1453
          |||         |||:::      :::
Db   4087 GGAACTCCTTCTCCCTCTCAGGGGAATGAG 4116
```

RESULT 12
AAV49652
ID   AAV49652 standard; DNA; 5802 BP.
XX
AC   AAV49652;
XX
DT   23-OCT-1998 (first entry)
XX
DE   Human SRCR protein DNA.

```
XX
KW   Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
KW   nervous system; medullo-blastoma; glioma; breast; detection;
KW   autoantibody; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             107..5464
FT                   /*tag= a
FT                   /product= SRCR protein"
FT                   /note= "Scavenger receptor cysteine-rich domain
FT                          containing protein"
XX
PN   WO9830687-A2.
XX
PD   16-JUL-1998.
XX
PF   09-JAN-1998;   98WO-DE00096.
XX
PR   18-JUL-1997;   97DE-1030997.
PR   09-JAN-1997;   97DE-1000519.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI   Mollenhauer J,  Poustka A;
XX
DR   WPI; 1998-399136/34.
DR   P-PSDB; AAW64591.
XX
PT   Proteins containing scavenger receptor, cysteine rich domain -
PT   useful for diagnosis and treatment of tumours
XX
PS   Claim 4; Fig 2; 54pp; German.
XX
CC   This sequence encodes a human protein which contains a SRCR (scavenger
CC   receptor, cysteine-rich) domain. The gene and encoded protein can be used
CC   to diagnose or treat tumours, particularly of the nervous system
CC   (medullo-blastoma or glioma) or breast. The DNA sequence and probes
CC   derived from it, are used to identify genes that express SRCR-domain
CC   containing proteins, to determine the form in which these proteins exist
CC   and to assess the significance of individual forms on cellular
CC   properties. The protein can be used to detect the presence of
CC   autoantibodies and antibodies which regulate its expression.
XX
SQ   Sequence 5802 BP; 1253 A; 1614 C; 1576 G; 1359 T; 0 other;

Alignment Scores:
Pred. No.:              1.37e-146     Length:        5802
Score:                  2435.00       Matches:       533
Percent Similarity:     46.58%        Conservative:  155
Best Local Similarity:  36.09%        Mismatches:    415
Query Match:            29.92%        Indels:        374
DB:                     19            Gaps:          28

US-09-759-130B-381 (1-1453) x AAV49652 (1-5802)

Qy      43 GlyThrAsp------LeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
              |||:::|||      |||   |||||||||||||||||||||||    |||    |||
Db     389 GGATCTGATTCTGGTTTGGCCCTGAGGCTGGTGAATGGAGATGGCAGGTGTCAGGGCCGA 448

Qy      61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
           ||||||:::   :::::::|||   |||||||||||||||||||||    |||:::|||
Db     449 GTGGAGATCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGACACCAAT 508

Qy      81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMet--------- 97
              :::    |||||||||:::||||||||||||     ::::::    :::
Db     509 GATGCCAACGTGGTCTGTAGGCAGCTGGGTTGTGGCTGGGCCATGTCAGCTCCAGGAAAT 568

Qy      98 PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCys 117
                 |||||||||               |||   |||   ||||||||||||    |||
Db     569 GCCTGGTTTGGCCAGGGC---------TCAGGACCCATTGCCCTGGATGATGTGCGCTGC 619

Qy     118 TyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCys 137
              |||:::||||||    ||||||   |||    |||       |||    |||||||||||
Db     620 TCAGGACACGAATCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGT 679

Qy     138 TyrHisGlyGluAspValGlyValAsnCysTyr------------------------- 148
              ||||||||||||    ||||||   |||
Db     680 GGCCATGGTGAAGATGCTGGTGTTATCTGCTCAGCTGCCCAGCCTCAGTCAACACTCAGG 739

Qy     149 ---------------------------------------------GlyGluAlaAsn 152
                                                         |||::::::
Db     740 CCAGAAAGTTGGCCTGTCAGGATATCACCACCTGTACCCACAGAAGGATCTGAATCCAGT 799

Qy     153 LeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPhe 172
           |||    ||||||||||||:::|||    :::    |||    ||||||||||||||    :::
Db     800 TTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGGTCCTATAC 859

Qy     173 GlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValVal 192
           :::      |||||||||:::|||||||||    |||:::    |||    |||    ||||||
Db     860 CGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGTC 919

Qy     193 CysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla 212
           |||||||||||||||||          :::       :::   |||
Db     920 TGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAG 979

Qy     213 ValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrp 232
                    ||||||   |||||||||:::   |||    |||:::|||        ||||||
Db     980 GGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCCTACCTGTGG 1039

Qy     233 AsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeu 252
           :::|||   |||   ||||||         |||:::|||   |||:::||||||       :::
Db    1040 AGCTGCCCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGTGAAGACGCTGGTGTC 1099

Qy     253 ThrCysTyr------------------------------------------------- 255
              |||
Db    1100 ATCTGCTCAGCTCCCCAGTCCCGGCCGACACCCAGCCCAGATACTTGGCCGACCTCACAT 1159

Qy     256 ---------------AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArg 270
                          |||   |||   ||||||||||||     |||    :::|||
Db    1160 GCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGG 1219

Qy     271 CysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLys 290
           |||    ||||||||||||:::          :::|||   ||||||||||||||
Db    1220 TGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGATAGC 1279

Qy     291 TrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHis 310
           |||:::    :::    |||:::|||||||||:::||||||||||||||    |||
Db    1280 TGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGTGGCTGGGCCACGTCA 1339

Qy     311 PheAlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSer 330
                 |||                  ||||||     :::    ||||||    |||
Db    1340 GCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTGGATGACGTGCGC 1399

Qy     331 CysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAsp 350
           |||||||||    ||||||:::||||||   |||    |||:::|||    ::::::    :::
Db    1400 TGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAAC 1459

Qy     351 CysLeuHisGlnAsnAspValSerValIleCysSerAspGlyAlaAspLeuGluLeuArg 370
           |||    |||       |||          |||||||||||
Db    1460 TGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCC----------------- 1501

Qy     371 LeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrp 390
                                                                     |||
Db    1502 -----------------------------------------------CACTCCTGG 1510

Qy     391 TrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu 410
              |||
Db    1511 TCGACG--------------------------------------------------- 1516

Qy     411 GlyCysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAsp 430
                                                      ||||||    :::
Db    1517 --------------------------------------CCCAGTCCAGACACATTGCCG 1537

Qy     431 IleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyr 450
```

```
Db   1538 ACCATCACCTTGCCTGCATCGACAGTAGGATCTGAATCCAGT------------------ 1579

Qy    451 AspGlyLysAlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSerAsp 470

Db   1579 ------------------------------------------------------------ 1579

Qy    471 LysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGlu 490

Db   1580 ---------TTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAG 1630

Qy    491 ValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsnAla 510

Db   1631 GTCCTATACCAAGGCTCCTGGGGCACCGTGTGCGATGACAGCTGGGACACCAATGATGCC 1690

Qy    511 AlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThrTyr 530

Db   1691 AATGTCGTCTGCAGGCAACCGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCGG 1750

Qy    531 PheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSer 550

Db   1751 TTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCT 1810

Qy    551 AsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAsp 570

Db   1811 TACCCGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATAGTGAAGAC 1870

Qy    571 ValIleValThrCysSerGlyAspAla------------------------ThrTrp--- 581

Db   1871 GCTGGTGTCATCTGCTCAGCTTCCCAGTCCCGGCCAACACCTAGTCCAGACACTTGGCCA 1930

Qy    582 ---------------------------------------GlyLeuArgLeuValGlyGly 588

Db   1931 ACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGA 1990

Qy    589 SerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrValCys 608

Db   1991 GGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGT 2050

Qy    609 AspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAspCysProSer 628

Db   2051 GATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGCTGG 2110

Qy    629 SerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAsp 648

Db   2111 GCCATGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTGGAT 2170

Qy    649 AspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGly 668

Db   2171 GATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGGCTC 2230

Qy    669 AsnAsnAspCysSerHisSerGluAspValGlyValIleCysSer--------------- 683

Db   2231 TCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCAG 2290

Qy    684 ---------------------------------------------------AspAlaSer 686

Db   2291 CCGACACCCAGCCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCTGAATCC 2350

Qy    687 AspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValGluValAsn 706

Db   2351 AGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTA 2410

Qy    707 ValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluVal 726

Db   2411 TACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTG 2470

Qy    727 ValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluProHisPheThr 746

Db   2471 GTTTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGC 2530

Qy    747 GluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeu 766

Db   2531 CAGGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTATCTG 2590

Qy    767 TrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSer 786

Db   2591 TGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCTGGT 2650

Qy    787 LeuIleCysSerAlaHisArg---GlnProArgLeuValGlyAlaAspMetProCysSer 805

Db   2651 GTCATCTGCTCAGCTTCCCAGTCCCAGCCG------------------------------ 2680

Qy    806 GlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSer 825

Db   2681 -----------ACACCCAGCCCAGACACTTGGCCAACC---------------------- 2707

Qy    826 LeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSer 845

Db   2707 ------------------------------------------------------------ 2707

Qy    846 ValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGlu 865

Db   2708 ------------------------------------------TCTCGTGCATCAACAGCA 2725

Qy    866 GlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIle 885

Db   2726 GGATCTGAATCCACTTTGGCCCTG------------------------------------ 2749

Qy    886 HisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGly 905

Db   2750 ---------------------------------------------AGACTGGTGAATGGA 2764

Qy    906 LysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCys 925

Db   2765 GGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGT 2824

Qy    926 AspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThr 945

Db   2825 GATGACTACTGGGACACCAATGATGCCAACGTGGTCTGCAGGCAGCTGGGCTGTGGCTGG 2884

Qy    946 AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHis 965

Db   2885 GCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGAT 2944

Qy    966 ArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGly 985

Db   2945 GATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTC 3004

Qy    986 AlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGln 1005

Db   3005 TCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCTCAGTCCCAG 3064

Qy   1006 ------ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValPro 1023

Db   3065 TCAACGCCCAGGCCAGATACTTGGCTGACCACCAACTTACCGGCATTGACAGTAGGATCT 3124

Qy   1024 GluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSer 1043

Db   3125 GAATCCAGTTTG---------------------GCTCTGAGGCTGGTGAATGGAGGTGAC 3163

Qy   1044 ArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAsp 1063

Db   3164 AGGTGTCGAGGCCGAGTGGAGGTCCTGTATCGAGGCTCCTGGGGAACCGTGTGTGATGAC 3223

Qy   1064 GlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPhe 1083

Db   3224 AGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATG 3283

Qy   1084 AsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeu 1103

Db   3284 TCGGCCCCAGGAAATGCCCGGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGATGATGTG 3343

Qy   1104 AsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHis 1123

Db   3344 CGCTGCTCAGGGAATGAGTCCTACCTGTGGAGCTGCCCCCACAAAGGCTGGCTCACCCAC 3403
```

```
Qy   1124 AspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeu 1143
          :::|||   |||   |||||||||||||||||||||||          |||   ::::::
Db   3404 AACTGTGGCCATCACGAAGACGCTGGTGTCATCTGCTCA---------GCCACCCAAATA 3454

Qy   1144 TyrSerGluThr---------------------------------------GluThrGlu 1150
             |||   |||                                        :::
Db   3455 AATTCTACTACGACAGATTGGTGGCATCCAACAACTACAACCACTGCAAGACCCTCTTCA 3514

Qy   1151 SerCysAlaGlyArgLeuGluValPheTyr---AsnGlyThrTrpGlySer--------- 1166
          :::|||   |||   |||      ||||||   :::||||||:::   |||
Db   3515 AATTGTGGTGGCTTCTTA------TTCTATGCCAGTGGGACATTCTCCAGCCCATCCTAC 3568

Qy   1166 ------------------------------------------------------------ 1166

Db   3569 CCTGCATACTACCCCAACAATGCTAAGTGTGTTTGGGAAATAGAAGTGAATTCTGGTTAT 3628

Qy   1167 ---------ValGlyArgArgAsnIle--------------------------------- 1172
                   :::|||   |||:::
Db   3629 CGCATAAACCTGGGCTTCAGTAATCTGAAATTGGAGGCACACCATAACTGCAGTTTTGAT 3688

Qy   1173 ------------------------------ThrThrAlaIleAlaGlyIleValCys--- 1181
                                        ::::::   :::   |||   :::|||
Db   3689 TATGTTGAAATCTTTGATGGATCATTGAATAGCAGTCTCCTGCTGGGGAAAATCTGTAAT 3748

Qy   1182 ------ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu---------AlaPro 1196
                |||||||:::                  :::::::::            :::
Db   3749 GATACCAGGCAAATATTTACATCTTCTTACAACCGAATGACCATTCACTTTCGAAGTGAC 3808

Qy   1197 LeuSerLysThrGlySerGlyPheMet---TrpValAspAspIleGlnCysProLysThr 1215
          :::|||            :::|||||||:::   |||   :::
Db   3809 ATCAGTTTCCAAAACACTGGCTTTTTGGCTTGGTATAAC--------------------- 3847

Qy   1216 HisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAla 1235
                                                           |||   |||:::
Db   3848 ------------------------------------------------TCCTTCCCAAGC 3859

Qy   1236 GluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCys 1255
          :::   |||                       :::   :::                |||
Db   3860 GATGCCACCTTG-----------AGGTTGGTCAATTTAAATTCATCCTATGGTCTATGT 3907

Qy   1256 SerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrp 1275
          :::||||||||||||||||:::|||   |||:::|||||||||||||||||||||||||||
Db   3908 GCCGGGCGTGTAGAAATTTACCATGGTGGCACCTGGGGGACAGTTTGTGATGACTCCTGG 3967

Qy   1276 AspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAla 1295
             :::   ||||||||||||||||||:::|||||||||||||||   |||::::::|||
Db   3968 ACCATTCAGGAAGCTGAGGTGGTCTGCAGACAGCTAGGGTGTGGACGTGCAGTTTCAGCC 4027

Qy   1296 LeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCys 1315
          |||   :::|||   ||||||   |||:::|||   |||   |||||||||:::   |||
Db   4028 CTTGGAAATGCATATTTTGGCTCTGGCTCTGGCCCCATCACCCTGGACGATGTAGAGTGC 4087

Qy   1316 LysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCys 1335
             |||   ||||||   ||||||   |||      :::   |||      :::|||
Db   4088 TCAGGGACGGAATCCACTCTCTGGCAGTGCCGGAACCGAGGCTGGTTCTCCCACAACTGT 4147

Qy   1336 GlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSer 1352
             |||:::|||||||||||||||   |||||||||      |||   :::
Db   4148 AATCATCGTGAAGATGCTGGTGTCATCTGCTCAGGAAACCATCTATCGACA 4198
```

```
RESULT 13
AAS17590
ID   AAS17590 standard; cDNA; 2697 BP.
XX
AC   AAS17590;
XX
DT   26-FEB-2002  (first entry)
XX
DE   DNA encoding novel secreted protein #19.
XX
KW   Secreted protein; cytostatic; immunosuppressive; vulnerary; vaccine;
KW   antiinflammatory; neuroprotective; nephrotropic; cardiovascular;
KW   human; cancer; autoimmune disease; wound healing disorder; infection;
KW   haematopoietic disorder; inflammatory disorder; infertility;
KW   neurological disease; psychiatric disease; cardiovascular disease;
KW   respiratory disease; renal; gastrointestinal; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2697
FT                   /*tag=  a
FT                   /product= "Human secreted protein"
XX
PN   WO200179454-A1.
XX
PD   25-OCT-2001.
XX
PF   11-APR-2001; 2001WO-US11797.
XX
PR   13-APR-2000; 2000US-196603P.
PR   24-APR-2000; 2000US-199417P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Agarwal P,  Murdoch PR,  Rizvi SK,  Smith RF,  Xiang Z;
XX
DR   WPI; 2002-061975/08.
DR   P-PSDB; AAU09878.
XX
PT   New secreted proteins or polypeptides, useful for treating e.g. cancer,
PT   autoimmune diseases, wound healing disorder, infections, haematopoietic
PT   disorders, inflammatory disorders, infertility, cancer  -
XX
PS   Claim 2; Page 47-48; 92pp; English.
XX
CC   The invention relates to an isolated novel secreted polypeptide (I) and
CC   polynucleotide (II). (I) and (II) are useful for treating cancer,
CC   autoimmune diseases, wound healing disorder, infections, haematopoietic
CC   disorders, inflammatory disorders, infertility, neurological and
CC   psychiatric diseases, cardiovascular diseases, respiratory diseases,
CC   renal diseases, or gastrointestinal diseases. These may also be used to
CC   treat diseases, abnormalities and disorders caused by abnormal
CC   expression, production, function and/or metabolism of the genes, as
CC   vaccines for inducing immunological response in a mammal, and in
CC   screening methods for detecting the effect of added compounds on the
CC   production of mRNA and polypeptide in cells. The polypeptides can be used
CC   as immunogens to produce antibodies immunospecific for the polypeptides,
CC   and to identify membrane-bound or soluble receptors. The polynucleotides
CC   may be used as diagnostic reagents, in chromosome localisation studies,
CC   and in tissue expression studies. The present sequence represents the
CC   coding sequence of novel human secreted protein #19.
XX
SQ   Sequence 2697 BP; 358 A; 841 C; 1055 G; 443 T; 0 other;

Alignment Scores:
Pred. No.:              6.94e-104     Length:        2697
Score:                  1763.50       Matches:       361
Percent Similarity:     49.37%        Conservative:  108
Best Local Similarity:  38.00%        Mismatches:    262
Query Match:            21.67%        Indels:        219
DB:                     24            Gaps:          16

US-09-759-130B-381 (1-1453) x AAS17590 (1-2697)

Qy    536 GlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCys 555
          |||||||:::   |||                                   |||:::|||
Db     31 GGGCCCCTTCTTCTG-----------------------------AATCTCTGGGCAGTC 60

Qy    556 GluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCys 575
                   |||   |||
Db     61 CCCATTGGTGGACCAGGT------------------------------------------ 78

Qy    576 SerGlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArg 595
```

```
Db     79 -----------------GCTCTGAGGCTGGCGTACAGACACAGCACGTGCGACGGAGTG 120

Qy    596 LeuGluValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLys 615
Db    121 GTGTTGGTCCGACACCACGGGGCATGGGGATACGTGTGCAACCAGGAGTGGACGCTGGCA 180

Qy    616 AlaAlaAlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeu 635
Db    181 GAGGCCTCTGTCGTGTGCAGGCAGCTGGGCTGCGGCCCTGCCGTGGGCGCCCCCAAGTAT 240

Qy    636 GlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAsp 655
Db    241 GTCCCGCTGCCTGGAGAGATGGCCCAGCCCTGGCTTCACAACGTGTCCTGCCGGGGCAAC 300

Qy    656 GluSerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSer 675
Db    301 GAGTCCTCCCTCTGGGAGTGCAGCCTTGGCTCATGGTGCCAGAGCCCGTGCCCCCACGCA 360

Qy    676 GluAspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGly 695
Db    361 TGGGTGGTGGTCGCGCTGTGCTCCAACGGCACTTTCCGGGAGCTCCGGCTGGTGAAGGGC 420

Qy    696 SerSerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCys 715
Db    421 CGCAGTCCCTGCGCGGGACTCCCCGAGATCAGAAACGTGAATGGGGTGGACCGCCTCTGT 480

Qy    716 AlaAsnGlyTrpGlyMetAsnIleAlaGlu------ValValCysArgGlnLeuGluCys 733
Db    481 GTC-----------CTGCATGTGGAGGAGGCCATGGTGTTCTGCCGGGAGCTGGGGTGC 528

Qy    734 GlySerAlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeu 753
Db    529 GGCCCTGTGCTCCAGGCCCCCCGC-----------------CGGGACGTGGGCGTCGTC 570

Qy    754 MetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGlu 773
Db    571 AGGAAGTACCTGGCCTGCAGGGGTACCGAGCCCACCATCCGCAGCTGCAGACTGGACAAC 630

Qy    774 TrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArg 793
Db    631 AACTTCCGCAGCGGCTGCGACCTGCGGCTGGACGCAGAGGTGGTCTGCTCAGGACACACC 690

Qy    794 GlnProArgLeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAla 813
Db    691 GAGGCCCGACTGGTGGGCGGCGAGCACCCCTGCGCCGGGCGCCTGGAGGTG--------- 741

Qy    814 AspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCys 833
Db    742 ---ACCTGGGGCACCGTCTGTGATGCGGCCCTGGACCTGGCCACGGCCCACGTGGTGTGC 798

Qy    834 ArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly 853
Db    799 CGGGAGCTGCAGTGTGGGGCGGTCGTGTCCACGCCCGAGGGCGCCCGCTTCGGCCGGGGC 858

Qy    854 AsnGlyLeuThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeu 873
Db    859 TCGGGGCCGGTGTGGACGGAGGCCTTCCGCTGTGCGGGCAACGAGTCGCTGCTGTTCCAC 918

Qy    874 CysProIleValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValVal 893
Db    919 TGCCCA------CGGGGGCGTGGGAGCCAGTGTGGGCATGGTCACGACGCGGGGCTCAGG 972

Qy    894 CysSerArgTyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnVal 913
Db    973 TGCTCAGAGTTC---------AGGATGGTCAACGGCAGCAGCAGCTGTGAGGGCCGCGTG 1023

Qy    914 GluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAsp 933
Db   1024 GAGTTCCAGGTGCAGGGGTCCTGGGCACCCCTCTGTGCCACCCACTGGGACATAGCAGAT 1083

Qy    934 AlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLys 953
Db   1084 GCCACCGTCCTCTGCCACCAGCTCAACTGTGGCAACGCGGTGGCCGCACCTGGAGGAGGC 1143

Qy    954 TyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGlu 973
Db   1144 CATTTTGGGGACGGGGACGCTGCCATCTGGCCTGATGCCTTTCACTGTGAGGGGACAGAG 1203

Qy    974 SerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsn 993
Db   1204 TCCTACTTGTGGAATTGCCCAGTAAGCACCCTGGGGGCCCCGGCCTGTGCCCCGGGAAAC 1263

Qy    994 ThrValSerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsn 1013
Db   1264 ACAGCCTCCGCGGTCTGCTCAGGT-------------------------CTGGCCCAC 1296

Qy   1014 ValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAsp 1033
Db   1297 GCC--------------------------------------------------------- 1299

Qy   1034 LysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIle------ 1051
Db   1300 ------CTGCGACTGAGGGAAGGACAGAGCCGCTGTGACGGCCGCGTGGAGGTCTCCCTG 1353

Qy   1051 ------------------------------------------------------------ 1051
Db   1354 GATGGCGTGTGGGGCCGCGTCCTGGACGATGCCTGGGACCTGCGCGGCGCGGGCGTGGTG 1413

Qy   1051 ------------------------------------------------------------ 1051
Db   1414 TGCCGGCAACTCGGGTGCAGAGGGGCCCAGCAAGCCTATGACGCACCTGCCCCCAGCCGC 1473

Qy   1051 ------------------------------------------------------------ 1051
Db   1474 GGATCCGTCCAGGTGGCGCTGAGCCGCGTGCGCTGTCTGGGCACCGAAACCCGCCTGACT 1533

Qy   1051 ------------------------------------------------------------ 1051
Db   1534 CAGTGCAACGTGTCCGCGACCCTGCAGGAGCCCGCGGGGACCTCGCGGGACGCCGGCGTG 1593

Qy   1051 ------------------------------------------------------------ 1051
Db   1594 GTGTGCTCCGGTGAGGTCGGAACCGCGTCCCCCATGGCCCGTCGCCACGGGATCCCGGGC 1653

Qy   1051 ------------------------------------------------------------ 1051
Db   1654 GCCCTGACTCTGTCTCTCCACAGGGAGCCTCAGGGTGCGGCTGGCCGCGGGGCCGGGGCG 1713

Qy   1052 TyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHis 1071
Db   1714 CTGCACGGGGGCGCGTGGGGCACCGTGTGTGACGATGCCTGGGACCTGCGGGACGCGCAC 1773

Qy   1072 ValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPhe 1091
Db   1774 GTGGTCTGCAGGCAGCTGGGCTGTGGCCGCGCCCTGAGCGCCCTGGGGGCCGCACACTTC 1833

Qy   1092 GlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHis 1111
Db   1834 GGAGCCGGGGCAGGGCGCATCTGGCTGGACGAGCTGGGCTGCCAGGGCCACGAGTCTGCG 1893

Qy   1112 LeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAla 1131
Db   1894 CTGTGGCAGTGCCCGTCGGCGGGCTGGGGGCGGCACGACTGGAGGCACAAGGAGGACGCC 1953

Qy   1132 GlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151
Db   1954 GGCGTCTTCTGCTCAGAGTCGGTGGCTCTGAGGCTG------CGAGGTGGGACCTGCTGC 2007

Qy   1152 CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn 1171
Db   2008 TGTGCTGGGTGGCTGGACGTGTTCTACAATGGGACCTGGGGCGCCATGTGCAGCAATGCC 2067

Qy   1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191
Db   2068 CTGAAGGACCTCTCCTTGTCCATCATCTGCAAGCAGCTGGGGTGTGGGGTGTGGGGA--- 2124
```

```
Qy    1192 ValSerLeuAla---------ProLeuSerLysThrGlySerGlyPheMetTrpValAsp 1208
            |||  ||||||           |||         |||:::|||     ||||||||
Db    2125 GTGGGGCTGGCTGGAGAACAGGCCCTTCCCCTCGCGGGCACCGGGACCGCCTGGGTGGAC 2184

Qy    1209 AspIleGlnCysProLys---ThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp 1227
            :::|||:::|||   :::      :::   ::::::||||||||   |||   ||||||
Db    2185 AACATCGAGTGCCGCAGGCTGCCCAACTCCACTCTGTGGCAATGCCCTTCCCACCCATGG 2244

Qy    1228 GluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys--------------- 1242
                                    |||:::   ||||||||||||
Db    2245 CACCCGCACTCTTGCGACCTTCGAGAGCAGGTCTGGATTACCTGTGCAGTGACCGCAGCC 2304

Qy    1243 ------------GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArg 1258
                        |||    :::|||||||||||||||:::        ||||||||||||
Db    2305 CCTTTTGCAGAGGAGGGCGCACTGCGCGTGCGCGGGGGCGAGGACCGCTGCTCCGGGCGC 2364

Qy    1259 ValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAla 1278
            ||||||:::||||||||||||||||||||||||||||||||||||   ||||||||||||
Db    2365 GTGGAGCTCTGGCACGCGGGCTCCTGGGGCACCGTGTGCGACGATGGCTGGGACCTGGCG 2424

Qy    1279 GluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAsp 1298
            :::|||||||||||||||:::||||||||||||||||   |||:::|||||||||
Db    2425 GACGCGGAGGTCGTGTGCCGCCAGCTGGGCTGTGGTCGGGCCGTCGCCGCCCTGGGGGCC 2484

Qy    1299 AlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsn 1318
            |||:::||||||   |||:::|||   :::||||||||||||::::::   |||:::|||:::
Db    2485 GCCGCCTTTGGCCCTGGCTCCGGGCCCGTGTGGCTGGACGAGGTGGGGTGCCGGGGCAGC 2544

Qy    1319 GluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLys 1338
            |||:::   ||||||   |||   |||:::   ||||||:::   |||       |||:::
Db    2545 GAGGCGTCCCTGTGGGGCTGCCCTGCGGAGCGGTGGGGACGCGGAGACCGCGCGCACGAG 2604

Qy    1339 GluAspAlaGlyValArgCysSerGlyGln 1348
            |||||||||||||||||||||   |||:::
Db    2605 GAGGACGCGGGCGTGCGCTGCTGGGGTGAG 2634
```

RESULT 14
AAA43017

```
ID   AAA43017 standard; cDNA; 820 BP.
XX
AC   AAA43017;
XX
DT   21-AUG-2000  (first entry)
XX
DE   Human secreted expressed sequence tag SEQ ID NO:1757.
XX
KW   Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
KW   expressed sequence tag; EST; probe; chemotactic; proliferative;
KW   immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW   thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW   antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW   antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW   cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW   vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW   insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW   lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW   central nervous system disorder; Alzheimer's disease; stroke;
KW   Parkinson's disease; Huntington's disease; coagulation disorder;
KW   haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW   tumour; infection; depression; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200021990-A1.
XX
PD   20-APR-2000.
XX
PF   15-OCT-1999;   99WO-US24205.
XX
PR   15-OCT-1998;   98US-0104435.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M;
XX
DR   WPI; 2000-317937/27.
XX
PT   Isolated polynucleotides, and encoded proteins, comprising secreted
PT   expressed sequence tags (sESTs), useful for treating various disorders
PT   such as autoimmune, infectious, and central nervous system disorders -
XX
PS   Claim 1; Page 527; 618pp; English.
XX
CC   AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC   sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC   tissue sources. The sESTs can have a range of activities depending on
CC   the tissues they were isolated from. The activities include:
CC   chemotactic; proliferative; immunomodulatory; haematopoietic;
CC   chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC   cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC   antiasthmatic; vulnerary; antiulcer; osteopathic; neuroprotective;
CC   nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC   anticonvulsant; and antidepressant. The sESTs can be used for gene
CC   therapy and in vaccines. The sESTs are useful as probes for the
CC   identification and isolation of full-length cDNAs and genomic DNA
CC   molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC   are useful in assays for determining biological activity and raising
CC   antibodies. They may be useful for treatment of autoimmune disorders
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression and
CC   psoriasis. AAA43420 to AAA43425 represent linker variants which are given
CC   in the exemplification of the present invention.
XX
SQ   Sequence 820 BP; 195 A; 155 C; 233 G; 236 T; 1 other;

Alignment Scores:
Pred. No.:              1.94e-81      Length:         820
Score:                  1404.00       Matches:        253
Percent Similarity:     96.98%        Conservative:   4
Best Local Similarity:  95.47%        Mismatches:     7
Query Match:            17.25%        Indels:         2
DB:                     21            Gaps:           0

US-09-759-130B-381 (1-1453) x AAA43017 (1-820)

Qy       1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHisGlnAsn 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26 ATGATGCTGCCTCAAAACTCGTGGCATATTGATTTTGGAAGATGCTGCTGTCATCAGAAC 85

Qy      21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 CTTTTCTCTGCTGTGGTAACTTGCATCCTGCTCCTGAATTCCTGCTTTCTCATCAGCAGT 145

Qy      41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 205

Qy      61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrp-AsnThrTh 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     206 GTGGAGGTGAAATTCCAGGGACAGTGGGGGACTGTGTGTGATGATGGGTGGGAACACTAC 265

Qy      80 rAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPh 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     266 TGCCTCAACTGTCGTGTGCAAACAGCTTGGATGTCCATTTTCTTTCGCCATGTTTCGTTT 325

Qy     100 eGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAs 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     326 TGGACAAGCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCCTGTTATGGAAA 385

Qy     120 nGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGl 140
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      386 TGAGTCAGCTCTCTGGGAATGTCAACACCGGGAATGGGGAAGCCATAACTGTTATCATGG 445

Qy      140 yGluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGl 160
             ||||:::||||||||||||||||         ::::::||||||||||||||||||||||
Db      446 AGAAGAAGTTGGTGTGAACTGTTAACGGTGAAGCCATCTGGGTTTGAGGCTAGTGGATGG 505

Qy      160 yAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCy 180
             ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      506 AAACAC-TCCTGTTCAGGGAGAGTGGAGGTGAAATTCCAAGAAAGGTGGGGAACTATATG 564

Qy      180 sAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSe 200
             |||||||||||||||||||||||     ||||||:::|||||||||||||||||||||||
Db      565 TGATGATGGGTGGAACTTAAATACCCCTGCCGTCCTGTGCAGGCAACTAGGATGTCCATC 624

Qy      200 rSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAs 220
             |||||||||||||||||||||     ||||||||||||||||||||||||||||||||||
Db      625 TTCTTTTATTTCTTCTGGAGTTGCTAACAGCCCTGCTGTATTGCGCCCCATTTGGCTGGA 684

Qy      220 pAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGl 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      685 TGACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTGCAGACATCGTGGATGGGG 744

Qy      240 yAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLe 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      745 AAATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTTATGATAGTAGTGATCT 804

Qy      260 uGluLeuArgLeu 264
             ||||   ||||||
Db      805 TGAACGTAGGCTC 817
```

RESULT 15
AAH99258

```
ID   AAH99258 standard; cDNA; 690 BP.
XX
AC   AAH99258;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein encoding cDNA sequence SEQ ID NO:93.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US35017.
XX
PR   23-DEC-1999;   99US-0471275.
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   P-PSDB; AAM25317.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS   Claim 1; Page 341; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and
CC   cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production. The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders.
XX
SQ   Sequence 690 BP; 163 A; 134 C; 207 G; 186 T; 0 other;

Alignment Scores:
Pred. No.:               2.42e-67      Length:        690
Score:                   1183.00       Matches:       214
Percent Similarity:      99.07%        Conservative:  0
Best Local Similarity:   99.07%        Mismatches:    2
Query Match:             14.54%        Indels:        0
DB:                      22            Gaps:          0

US-09-759-130B-381 (1-1453) x AAH99258 (1-690)

Qy      636 GlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAsp 655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       43 GGAAACGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTGATGGAGAT 102

Qy      656 GluSerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSer 675
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      103 GAGTCAGATCTCTGGTCATGCAGGAACAGTGGGTGGGGAAATAATGACTGCAGTCACAGT 162

Qy      676 GluAspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGly 695
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      163 GAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGA 222

Qy      696 SerSerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCys 715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      223 AGCAGCAGGTGTGCTGGAAAAGTTGAGGTGAATGTCCAGGGTGCCGTGGGAATTCTGTGT 282

Qy      716 AlaAsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySer 735
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      283 GCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGGTCT 342

Qy      736 AlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSer 755
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      343 GCAATCAGGGTCTCCAGAGAGCCTCATTTCACAGAAAGAACATTACACATCTTAATGTCC 402

Qy      756 AsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLys 775
            |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db      403 AATTCTGGCTGCGCTGGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAA 462

Qy      776 GlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnPro 795
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      463 CAGACTGCGTGTCATTTAAATATGGAAGCAAGTTTGATCTGCTCAGCCCACAGGCAGCCC 522
```

```
Qy    796 ArgLeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThr 815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db    523 AGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTTGAAGTGAAACATGCACACACA 582

Qy    816 TrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGlu 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 TGGCGCTCTGTCTGTGATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAA 642

Qy    836 LeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGly 851
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    643 TTAAACTGTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGGG 690
```

Search completed: May 12, 2003, 06:58:24
Job time : 670 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         May 12, 2003, 10:35:25 ; Search time 2862 Seconds
                                          (without alignments)
                                          8222.237 Million cell updates/sec

Title:          US-09-759-130B-381
Perfect score:  8138
Sequence:       1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09759130/runat_06052003_123554_23365/app_query.fasta_1.1607
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759130_@CGN_1_1_2168_@runat_06052003_123554_23365 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  gb_gss:*
                18:  em_gss_hum:*
                19:  em_gss_inv:*
                20:  em_gss_pln:*
                21:  em_gss_vrt:*
                22:  em_gss_fun:*
                23:  em_gss_mam:*
                24:  em_gss_mus:*
                25:  em_gss_other:*
                26:  em_gss_pro:*
                27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1765.5	21.7	999	9	AL529439	AL529439 AL529439
	2	1083	13.3	940	9	AL550994	AL550994 AL550994
	3	1072	13.2	582	12	BG570972	BG570972 602591546
	4	960	11.8	799	13	BI185454	BI185454 UNL-P-FN-
	5	942.5	11.6	1108	9	AL544534	AL544534 AL544534
	6	936.5	11.5	1100	9	AL540905	AL540905 AL540905
	7	929	11.4	718	13	BI183796	BI183796 UNL-P-FN-
c	8	895	11.0	554	9	AL529438	AL529438 AL529438
	9	836.5	10.3	930	13	BI838644	BI838644 603086185
c	10	767	9.4	1153	9	AL540904	AL540904 AL540904
	11	739	9.1	723	13	BI836459	BI836459 603082833
	12	718	8.8	644	13	BI009838	BI009838 MR4-RT004
	13	701	8.6	585	14	BQ327934	BQ327934 MR4-RT004
	14	695	8.5	653	10	BE502724	BE502724 hz78d12.x
c	15	679	8.3	616	13	BI009795	BI009795 MR4-RT004
	16	675	8.3	652	10	BB622216	BB622216 BB622216
	17	669.5	8.2	913	14	BQ710454	BQ710454 AGENCOURT
c	18	667.5	8.2	711	14	BQ378694	BQ378694 MR3-UT009
	19	667	8.2	574	14	BQ327932	BQ327932 MR4-RT004
c	20	666	8.2	581	14	BQ776070	BQ776070 UI-H-FH0-
	21	660.5	8.1	593	14	BQ327876	BQ327876 MR4-RT004
	22	658	8.1	574	13	BI009286	BI009286 MR4-RT004
	23	657	8.1	583	14	BQ327909	BQ327909 MR4-RT004
c	24	651	8.0	596	13	BI009802	BI009802 MR4-RT004
c	25	649	8.0	934	14	BQ936228	BQ936228 AGENCOURT
c	26	645.5	7.9	927	9	AL570629	AL570629 AL570629
c	27	634.5	7.8	689	12	BF096204	BF096204 MR3-UT009
	28	634	7.8	2925	11	AK014331	AK014331 Mus muscu
c	29	626	7.7	505	13	BI038644	BI038644 IL3-NT028
	30	626	7.7	705	17	B16308	B16308 34707.TP CI
c	31	625	7.7	607	14	BQ327933	BQ327933 MR4-RT004
	32	624	7.7	3665	11	BC035470	BC035470 Homo sapi
	33	621	7.6	523	12	BG534729	BG534729 602553892
	34	620	7.6	569	10	AW631641	AW631641 90867 MAR
c	35	616.5	7.6	833	9	AL574811	AL574811 AL574811
c	36	616	7.6	560	13	BI009799	BI009799 MR4-RT004
c	37	612	7.5	529	13	BI009105	BI009105 MR4-RT004
	38	612	7.5	531	14	BQ327873	BQ327873 MR4-RT004
c	39	612	7.5	579	14	BQ327906	BQ327906 MR4-RT004
	40	607	7.5	602	12	BG641637	BG641637 pg11c.pk0
c	41	605	7.4	583	14	BQ368412	BQ368412 PM3-GN051
c	42	587.5	7.2	594	10	AV616074	AV616074 AV616074
	43	579	7.1	822	12	BF792807	BF792807 602253274
	44	578	7.1	1581	11	BC020481	BC020481 Homo sapi
c	45	577.5	7.1	850	9	AL550308	AL550308 AL550308

ALIGNMENTS

```
RESULT 1
AL529439
LOCUS       AL529439                999 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL529439 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD006YO05 5
            prime, mRNA sequence.
ACCESSION   AL529439
VERSION     AL529439.1  GI:12792932
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 999)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
```

```
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .999
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DD006YO05"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      220 a     236 c     318 g     224 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               5.88e-168      Length:          999
Score:                   1765.50        Matches:         327
Percent Similarity:      98.50%         Conservative:    1
Best Local Similarity:   98.20%         Mismatches:      5
Query Match:             21.69%         Indels:          3
DB:                      9              Gaps:            0

US-09-759-130B-381 (1-1453) x AL529439 (1-999)

Qy     1060 IleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCys 1079
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCCACGTGGTGTGTCAAAAGCTGGGCTGT 60

Qy     1080 GlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrp 1099
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GGAGTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTCAGGGCCCATCTGG 120

Qy     1100 LeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGly 1119
            ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db      121 CTGGATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGC 180

Qy     1120 TrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThr 1139
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TGGGGGCAGCACGACTGCAGGCACAAGGAGGACGCAGGGGTCATCTGCTCAGAATTCACA 240

Qy     1140 AlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPhe 1159
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GCCTTGAGGCTCTACAGTGAAACTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTC 300

Qy     1160 TyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIle 1179
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCACCACAGCCATAGCAGGCATT 360

Qy     1180 ValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLys 1199
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTAGTCAGCCTCGCCCCTTTATCTAAG 420

Qy     1200 ThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIle 1219
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 ACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCCTAAAACGCATATCTCCATA 480

Qy     1220 TrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrp 1239
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGAATCTCCAGCCCAGCAGAAGAGACCTGG 540

Qy     1240 IleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgVal 1259
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 ATCACATGTGAAGATAGAATAAGAGTGCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTG 600

Qy     1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GAGATCTGGCACGCAGGCTCCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGCCGAG 660

Qy     1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAla 1299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GCGGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCT 720

Qy     1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 TCGTTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGCGGTGCAAAGGAAATGAG 780

Qy     1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TCATTTCTATGGGACTGTCACGCCAAACCCTGGGGACAGAGTGACTGTGGACACAAGGAA 840

Qy     1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GATGCTGGCGTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCAT 900

Qy     1360 LeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPhe 1379
            ||||||||||||||||||||||||||||||     ||||||||||||||:::   ||||||
Db      901 TTAGCACTTATTTTATCCAGTATCTTTGGGCT-CTTCTCCTGGTTCTGTATAT-CTATTT 958

Qy     1380 LeuThrTrpCysArgValGlnLysGlnLysHisLeuPro 1392
            |||||||||||||||||||||      ||||||||||||
Db      959 CTCACGTGGTGCCGAGTTCAGAA-MAAAAACATCTGCCC 996

RESULT 2
AL550994
LOCUS       AL550994                940 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL550994 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI066YE07 5
            prime, mRNA sequence.
ACCESSION   AL550994
VERSION     AL550994.1  GI:12888516
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 940)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI066YE07"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      228 a     212 c     279 g     218 t      3 others
ORIGIN

Alignment Scores:
```

```
Pred. No.:              1.03e-98      Length:        940
Score:                  1083.00       Matches:       190
Percent Similarity:     75.24%        Conservative:  41
Best Local Similarity:  61.89%        Mismatches:    76
Query Match:            13.31%        Indels:        1
DB:                     9             Gaps:          0

US-09-759-130B-381 (1-1453) x AL550994 (1-940)

Qy    858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
          ||||||||:::||||||||||||||     |||:::|||||:::||||||    :::
Db      9 TGGGCTGAAGAATTCCAGTGTGAGGGACATGAGTCCCATCTTTCACTCTGCYCAGTAGCA 68

Qy    878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
                ||||||   ||||||   ||||||||:::||||||||||||||||||||||||||
Db     69 CCCCGCCCAGAAGGAACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATAC 128

Qy    898 ThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnVal 917
          |||::::::||||||||||||||||||:::     |||:::|||:::||||||:::
Db    129 ACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGTGAGGGCAGAGTGGAGCTCAAAACG 188

Qy    918 LeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeu 937
          ||||||   ||||||||||||||||:::::::||||||||||   |||||||||   ||||||
Db    189 CTTGGTGCCTGGGGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTT 248

Qy    938 CysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGlu 957
          |||:::||||||   ||||||   ||||||||||||   ||||||          |||:::
Db    249 TGCCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTTGGAAAA 308

Qy    958 ArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAsp 977
             :::   ::::::|||   |||   |||||||||   |||   |||       :::
Db    309 GGAAATGGTCAGATCTGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGA 368

Qy    978 AsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerVal 997
          :::|||   :::|||   |||||||||      |||             ||||||
Db    369 GATTGTCCTGTAACTGCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTA 428

Qy    998 IleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspPro 1017
          ||||||:::|||:::   :::|||   |||      |||   ::::::       |||
Db    429 ATCTGCTCAGGAAACCAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCA 488

Qy   1018 TyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArg 1037
                   :::||||||   |||||||:::   |||:::|||       :::||||||
Db    489 ACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGC 548

Qy   1038 LeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrp 1057
          ||||||:::|||      ||||||||||||||||||||||||||||:::|||     |||
Db    549 CTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGG 608

Qy   1058 GlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeu 1077
          ||||||||||||||||||   ||||||||||||||||||||||||||||::::::|||
Db    609 GGCACCATCTGTGATGACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTG 668

Qy   1078 GlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyPro 1097
          ||||||||   |||   |||||||||   ||||||||||||||||||||:::||||||
Db    669 GGCTGTGGAGAGGCCATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCC 728

Qy   1098 IleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSer 1117
          ||||||||||||::::::   |||   |||   ||||||   :::|||||||||   |||
Db    729 ATCTGGCTGGATGAGATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTGCCATTCA 788

Qy   1118 ArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGlu 1137
             ||||||||||||   :::|||||||||||||||||||||||||||||||||||||
Db    789 CACGGCTGGGGGCAGCAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAA 848

Qy   1138 PheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGlu 1157
          |||   :::|||||||||   ||||||         |||:::|||||||||   ||||||
Db    849 TTCATGTCTCTKAGACTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGGSGTCTGGAA 908

Qy   1158 ValPheTyrAsnGlyThrTrp 1164
             ||||||||||||   |||
Db    909 GT-TTTTACAATGGAGCTTGG 928
```

```
RESULT 3
BG570972
LOCUS       BG570972                 582 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602591546F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4714115 5',
            mRNA sequence.
ACCESSION   BG570972
VERSION     BG570972.1  GI:13578625
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1558  row: k  column: 12
            High quality sequence stop: 576.
FEATURES             Location/Qualifiers
     source          1..582
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4714115"
                     /clone_lib="NIH_MGC_79"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      126 a    140 c    198 g    118 t
ORIGIN

Alignment Scores:
Pred. No.:              5.38e-98      Length:        582
Score:                  1072.00       Matches:       192
Percent Similarity:     100.00%       Conservative:  1
Best Local Similarity:  99.48%        Mismatches:    0
Query Match:            13.17%        Indels:        0
DB:                     12            Gaps:          0

US-09-759-130B-381 (1-1453) x BG570972 (1-582)

Qy   1109 GluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLys 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 GAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCACGACTGCAGGCACAAG 61

Qy   1129 GluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGlu 1148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 GAGGACGCAGGGGTCATCTGCTCAGAATTCACAGCCTTGAGGCTCTACAGTGAAACTGAA 121

Qy   1149 ThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGly 1168
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 ACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGC 181

Qy   1169 ArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGlu 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 AGGAGGAACATCACCACAGCCATAGCAGGCATTGTGTGCAGGCAGCTGGGCTGTGGGGAG 241
```

```
Qy    1189 AsnGlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAsp 1208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 AATGGAGTTGTCAGCCTCGCCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGAT 301

Qy    1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGlu 1228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 GACATTCAGTGTCCTAAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAG 361

Qy    1229 ArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgVal 1248
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CGAAGAATCTCCAGCCCAGCAGAAGAGACCTGGATCACATGTGAAGATCGAATAAGAGTG 421

Qy    1249 ArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGly 1268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 CGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGGC 481

Qy    1269 ThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly 1288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 ACAGTGTGTGATGACTCCTGGGACCTGGCCGAGGCGGAAGTGGTGTGTCAGCAGCTGGGC 541

Qy    1289 CysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPhe 1301
           ||||||||||||||||||||||||||||||||||||:::
Db     542 TGTGGCTCTGCTCTGGCTGCCCTGAGGGACGCTTCGTAT 580
```

```
RESULT 4
BI185454
LOCUS       BI185454                799 bp    mRNA    linear   EST 10-JUL-2001
DEFINITION  UNL-P-FN-cw-d-01-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
            UNL-P-FN-cw-d-01-0-UNL 3', mRNA sequence.
ACCESSION   BI185454
VERSION     BI185454.1  GI:14659863
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Caetano,A.R., Johnson,R.K. and Pomp,D.
  TITLE     Generation and sequence characterization of a normalized cDNA
            library from swine ovarian follicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Pomp, D
            Department of Animal Science
            University of Nebraska, Lincoln
            Lincoln, NE 68583-0908, USA
            Tel: 402 472 6416
            Fax: 402 472 6362
            Email: dpomp@unl.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message.
            Seq primer: M13 -29
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..799
                     /organism="Sus scrofa"
                     /strain="University of Nebraska, Lincoln Swine Selection
                     Lines"
                     /db_xref="taxon:9823"
                     /clone="UNL-P-FN-cw-d-01-0-UNL"
                     /clone_lib="UNL-P-FN"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
                     library is a normalized library representing porcine
                     ovarian follicles, ranging between 2.0 to 10.0 mm in
                     diameter, collected  during 7 days of the follicular phase
                     of the pig estrous cycle. This  library was derived from
                     the library UNL-P-F2. The tag is a string of  5-6
                     nucleotides present between the Not I site and the
                     oligo-dT track.  The library was constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996.
                     TAG_SEQ=None found"
BASE COUNT      206 a    174 c    241 g    177 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               2.31e-86      Length:         799
Score:                   960.00        Matches:        165
Percent Similarity:      79.17%        Conservative:   44
Best Local Similarity:   62.50%        Mismatches:     55
Query Match:             11.80%        Indels:         1
DB:                      13            Gaps:           0

US-09-759-130B-381 (1-1453) x BI185454 (1-799)

Qy     999 CysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyr 1018
            |||:::|||:::   :::|||   |||   ||||||   ::::::   |||||||||
Db       9 TGCTCAGGGAACCAGAGTCAGACACTATCCCCGTGCAATTCATCATCCTCGGACCCATCA 68

Qy    1019 LeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeu 1038
              |||   :::   |||   :::   :::   |||:::            :::|||||||||
Db      69 AGCTCTATTATTTCAGAAGAAAATGGTGTTGCCTGCATAGGGAGTGGTCAACTTCGCCTG 128

Qy    1039 ValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGly 1058
           |||||||||   |||||||||||||||||||||||||:::|||||:::|||   ||||||
Db     129 GTCGATGGAGGTGGTCGTTGTGCTGGGAGAGTAGAGGTCTATCATGAGGGCTCCTGGGGC 188

Qy    1059 ThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGly 1078
           |||||||||||||||   ||||||||||:::|||||||||||||||||||||:::::::|||
Db     189 ACCATCTGTGATGACAGCTGGGACCTGAATGATGCCCATGTGGTGTGCAAACAGCTAAGC 248

Qy    1079 CysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIle 1098
           ||||||   |||   |||||||||   |||||||||||||||||||||:::|||||||||
Db     249 TGTGGATGGGCCATTAATGCCACTGGTTCTGCTCATTTTGGGGAAGGAACAGGGCCCATT 308

Qy    1099 TrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArg 1118
           |||||||||::::::||||||   |||   |||||||||:::|||||||||   |||
Db     309 TGGCTGGATGAGATAAACTGTAATGGAAAAGAATCTCATATTTGGCAATGCCACTCACAT 368

Qy    1119 GlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPhe 1138
           |||||||||:::|||:::|||||||||||||||||||||||||||||||||||||||||
Db     369 GGTTGGGGGCGGCACAATTGCAGGCATAAGGAGGATGCAGGAGTCATCTGCTCGGAGTTC 428

Qy    1139 ThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluVal 1158
              :::|||||||||   ||||||            |||:::||||||||||||||||||||
Db     429 ATGTCTCTGAGACTGATCAGTGAAAACAGCAGAGAGACCTGTGCAGGGCGCCTGGAAGTT 488

Qy    1159 PheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGly 1178
           ||||||||||||   |||||||||||||||:::   :::::::::   |||      |||
Db     489 TTTTACAACGGAGCTTGGGGCAGCGTTGGCAAGAATAGCATGTCTCCAGCCACAGTGGGG 548

Qy    1179 IleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSer 1198
           :::||||||||||||||||||||||   :::   |||   :::|||   |||
Db     549 GTGGTATGCAGGCAGCTAGGCTGTGCAGACAGAGGGGACATCAGCCCTGCATCTTCAGAC 608

Qy    1199 LysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSer 1218
           ||||||   |||   ||||||||||||:::::|||||||||||||            :::
Db     609 AAGACAGTGTCCAGGCACATGTGGGTGGACAATGTTCAGTGTCCTAAAGGACCTGACACA 668

Qy    1219 IleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThr 1238
           :::|||||||||   |||:::||||||::::::|||::::::||||||:::|||||||||
Db     669 CTATGGCAGTGCCCATCATCTCCATGGAAGAAGAGACTGGCCAGCCCCTCAGAGGAGACA 728

Qy    1239 TrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArg 1258
           ||||||||||||   ::::::||||||::::::   |||   |||   ||||||||||||
Db     729 TGGATCACATGTGCCAACAAAATAAGACTTCAAGAAGGAAN-ACTAATTGTTCTGGACGT 787

Qy    1259 ValGluIleTrp 1262
           ||||||||||||
Db     788 GTGGAGATCTGG 799

RESULT 5
AL544534
```

```
LOCUS       AL544534                1108 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL544534 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI021YH23 5
            prime, mRNA sequence.
ACCESSION   AL544534
VERSION     AL544534.1  GI:12877014
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1108)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..1108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI021YH23"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      307 a    189 c    319 g    289 t      4 others
ORIGIN

Alignment Scores:
Pred. No.:               2.58e-84       Length:          1108
Score:                   942.50         Matches:         169
Percent Similarity:      65.58%         Conservative:    52
Best Local Similarity:   50.15%         Mismatches:      111
Query Match:             11.58%         Indels:          5
DB:                      9              Gaps:            3

US-09-759-130B-381 (1-1453) x AL544534 (1-1108)

Qy       1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHis---Gln 19
            |||:::|||   :::::|||            ||||||  |||          |||
Db      69 ATGGTGCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGA--------CATTTTGTC 119

Qy      20 AsnLeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSer 39
            ||||||         :::|||   :::||||||||||:::::||||||:::   |||
Db     120 AACTTGAGTCCCTTCACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGT 179

Qy      40 SerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGly 59
            |||    |||||||||  ||||||||||||||:::|||:::       ||||||
Db     180 TCTCTTGGAGGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGNG 239

Qy      60 ThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThr 79
               ||||||||||||   |||    :::||||||||||||||||::::::||||||:::
Db     240 AGAGTGGAAGTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATG 299

Qy      80 ThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArg 99
               |||   :::|||:::|||   |||||||||||||||    :::
Db     300 GAAGCGGTCTCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGA 359

Qy     100 PheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGly 119
           :::      :::                |||:::||||||:::|||  |||||||||   |||
Db     360 TGGGCTAATTCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGG 419

Qy     120 AsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHis---AsnCysTyr 138
               ||||||||||||||||||:::|||:::|||        ||||||   |||   ||||||
Db     420 AATGAGTCAGCTCTTTGGGATTGCAAACATGATGGATGGGGAAAGCATAGTAACTGTACT 479

Qy     139 HisGlyGluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuVal 158
            |||   :::|||   ||||||   |||             :::||||||   :::||||||
Db     480 CACCAACAAGATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACG 539

Qy     159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
               |||    |||   ||||||||||||:::|||:::|||||||||   ||||||||||||
Db     540 CGTGGAGGGAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACA 599

Qy     179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
           :::||||||||||   :::|||:::::   |||:::|||:::||||||||||||||   |||
Db     600 GTGTGTGATGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGT 659

Qy     199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
              |||:::            ||||||       |||                   |||||||||
Db     660 GGAAGTGCTGTCAGTTTCTCTGGTTCATCTAATTTGGAGAAGGCTCTGGACCAATCTGG 719

Qy     219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
              ||||||::::::|||   |||||||||   ||||||||||||||||:::|||:::|||
Db     720 TTTGATGATCTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGA 779

Qy     239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSer 258
            ||||||  ||||::::|||    |||    ||||||      :::    |||       :::
Db     780 TGGGGAAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCA 839

Qy     259 AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLys 278
            ||||||   ||||||||||||   |||         |||   ||||||:::|||::::::
Db     840 GATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGA 899

Qy     279 IleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVal 298
              ||||||   |||||||||:::|||         |||::::::     |||   |||
Db     900 TTCCAAGGAGAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTG 959

Qy     299 ValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGln 318
              ||||||||||||||||||   |||||||:::          |||        :::
Db     960 GCATGCAAGCAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGCCATW 1019

Qy     319 SerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGlu 335
              |||         :::||||||||||   |||||||||   |||:::|||
Db    1020 VAGGGATTTGGACACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAA 1070

RESULT 6
AL540905
LOCUS       AL540905                1100 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL540905 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE004YI19 5 prime
            , mRNA sequence.
ACCESSION   AL540905
VERSION     AL540905.1  GI:12871463
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1100)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..1100
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DE004YI19"
                     /clone_lib="LTI_FL002_PL1"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
```

```
                  and cloned into the Not I and Eco RV sites of the
                  pCMVSPORT 6 vector.  Library was constructed by Life
                  Technologies. Contact : Feng Liang Life Technologies, a
                  division of Invitrogen 9800 Medical Center Drive Rockville
                  , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                  fliang@lifetech.com URL :
                  http://fulllength.invitrogen.com"
BASE COUNT      299 a    188 c    317 g    288 t      8 others
ORIGIN

Alignment Scores:
Pred. No.:              1.04e-83       Length:         1100
Score:                  936.50         Matches:        171
Percent Similarity:     65.40%         Conservative:   52
Best Local Similarity:  50.15%         Mismatches:     113
Query Match:            11.51%         Indels:         6
DB:                     9              Gaps:           3

US-09-759-130B-381 (1-1453) x AL540905 (1-1100)

Qy      1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysCysHis---Gln 19
          ||::::|||  :::::::|||             ||||||    |||       |||
Db     77 ATGGTGCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGA---------CATTTTGTC 127

Qy     20 AsnLeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSer 39
          |||||            :::|||  :::|||||||||::::::||||||:::    |||
Db    128 AACTTGAGTCCCTTCACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTTGTCACCAGT 187

Qy     40 SerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGly 59
          |||      |||||||||   ||||||||||||||||:::|||:::       |||||||||
Db    188 TCTCTTGGAGGAACAGACAAGGAGCTGAGGCTAGTGGATGGTGAAAACAAGTGTAGCGGG 247

Qy     60 ThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThr 79
             ||||||||||||    |||    :::|||||||||||||||::::::||||||:::
Db    248 AGAGTGGAAGTGAAAGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATG 307

Qy     80 ThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArg 99
             |||   :::|||:::|||   |||||||||||||||     :::
Db    308 GAAGCGGTCTCTGTGATTTGTAACCAGCTGGGATGTCCAACTGCTATCAAAGCCCCTGGA 367

Qy    100 PheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGly 119
          :::       :::               |||:::|||||||:::|||    |||||||||     |||
Db    368 TGGGCTAATTCCAGTGCAGGTTCTGGACGCATTTGGATGGATCATGTTTCTTGTCGTGGG 427

Qy    120 AsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHis---AsnCysTyr 138
          ||||||||||||||||||:::|||    |||        ||||||    |||    ||||||
Db    428 AATGAGTCAGCTCTTTGGGATTGCMAACATGATGGATGGGGAAAGCATAGTAACTGTACT 487

Qy    139 HisGlyGluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuVal 158
          |||    :::|||    ||||||    |||                :::||||||    :::||||||
Db    488 CACCAACAAGATGCTGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACG 547

Qy    159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
             |||    |||    |||||||||||||:::|||:::||||||||    ||||||||||||
Db    548 CGTGGAGGGAATATGTGTTCTGGAAGAATAGAGATCAAATTCCAAGGACGGTGGGGAACA 607

Qy    179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
          :::|||||||||    :::|||::::::    |||:::|||:::|||||||||||||    |||
Db    608 GTGTGTGATGATAACTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGT 667

Qy    199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
             |||:::          ||||||       |||                |||||||||
Db    668 GGAAGTGCTGTCAGTTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGG 727

Qy    219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
             ||||||::::::|||    |||||||||    |||||||||||||||:::|||:::|||
Db    728 TTTGATGATCTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAACATCAAGGA 787

Qy    239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSer 258
          ||||||    |||:::|||    |||    ||||||          :::    |||          :::
Db    788 TGGGGAAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCA 847

Qy    259 AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLys 278
             ||||||    ||||||||||||    |||          |||    ||||:::|||::::::
Db    848 GATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTCAGGAAGATTAGAAGTGAGA 907

Qy    279 IleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVal 298
             ||||||    |||||||||:::|||          |||::::::        |||    |||
Db    908 TTCCAAGGAGAATGGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTG 967

Qy    299 ValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGln 318
             ||||||||||||||||||    ||||||:::             |||          :::
Db    968 GCATGCAAGCAACTGGGATGTCCAACTGCCGTCACAGCCATTGGTCGAGTTAACGC-MRT 1026

Qy    319 SerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeu 338
             |||          :::|||||||||    |||||||||    |||:::|||          :::
Db   1027 WAGGGRTTTKGRCACATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTC 1086

Qy    339 Trp 339
          |||
Db   1087 TGG 1089

RESULT 7
BI183796
LOCUS       BI183796                 718 bp    mRNA    linear   EST 10-JUL-2001
DEFINITION  UNL-P-FN-cd-h-08-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
            UNL-P-FN-cd-h-08-0-UNL 3', mRNA sequence.
ACCESSION   BI183796
VERSION     BI183796.1  GI:14658205
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Caetano,A.R., Johnson,R.K. and Pomp,D.
  TITLE     Generation and sequence characterization of a normalized cDNA
            library from swine ovarian follicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Pomp, D
            Department of Animal Science
            University of Nebraska, Lincoln
            Lincoln, NE 68583-0908, USA
            Tel: 402 472 6416
            Fax: 402 472 6362
            Email: dpomp@unl.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message.
            Seq primer: M13 -29
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Sus scrofa"
                     /strain="University of Nebraska, Lincoln Swine Selection
                     Lines"
                     /db_xref="taxon:9823"
                     /clone="UNL-P-FN-cd-h-08-0-UNL"
                     /clone_lib="UNL-P-FN"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
                     library is a normalized library representing porcine
                     ovarian follicles, ranging between 2.0 to 10.0 mm in
                     diameter, collected  during 7 days of the follicular phase
                     of the pig estrous cycle. This  library was derived from
                     the library UNL-P-F2. The tag is a string of  5-6
                     nucleotides present between the Not I site and the
                     oligo-dT track.  The library was constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996.
                     TAG_SEQ=None found"
BASE COUNT      178 a    155 c    223 g    156 t      6 others
ORIGIN

Alignment Scores:
```

```
Pred. No.:               2.65e-83      Length:          718
Score:                   929.00        Matches:         156
Percent Similarity:      82.13%        Conservative:    37
Best Local Similarity:   66.38%        Mismatches:      42
Query Match:             11.42%        Indels:          1
DB:                      13            Gaps:            0

US-09-759-130B-381 (1-1453) x BI183796 (1-718)

Qy   1050 GluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069
            |||::::||||||:::|||   ||||||||||||||||||||   |||||||||:::|||
Db     14 GAGGTCTATCATGAGGGCTCCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAATGAT 73

Qy   1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089
          ||||||||||||||||:::::|||  ||||||   |||   |||||||||   ||||||
Db     74 GCCCATGTGGTGTGCAAACAGCTGAGCTGTGGATGGGCCATTAATGCCACTGGTTCTGCT 133

Qy   1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109
          ||||||||||||||||::::||||||||||||||||||||:::::::||||||   |||   |||
Db    134 CATTTTGGGGAAGGAACAGGGCCCATTTGGCTGGATGAGATAAACTGTAATGGAAAAGAA 193

Qy   1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129
          ||||||:::|||||||||   |||   |||||||||:::|||:::|||||||||||||||
Db    194 TCTCATATTTGGCAATGCCACTCACATGGTTGGGGGCGGCACAATTGCAGGCATAAGGAG 253

Qy   1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149
          ||||||||||||||||||||||||||||   :::|||||||||   ||||||
Db    254 GATGCAGGAGTCATCTGCTCGGAGTTCATGTCTCTCAGACTGATCAGTGAAAACAGCAGA 313

Qy   1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169
          |||:::||||||||||||||||||   |||||||||||||||   |||||||||||||||:::
Db    314 GAGACCTGTGCAGGGCGCCTGGNAGTTTTTTACAACGGAGCTTGGGGCAGCGTTGGCAAG 373

Qy   1170 ArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
             :::::::::   |||      |||:::||||||||||||||||||||||||   :::
Db    374 AATAGCATGTCTCCAGCCACAGTGGGGTGGTATGCAGGCAGCTGGGCTGTGCAGACAGA 433

Qy   1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
            |||   :::|||   |||         ||||||   |||      ||||||||||||:::
Db    434 GGGGACATCAGCCCTGCATCTTCAGACAAGACAGTGTCCAGGCACATGTGGGTGGACAAT 493

Qy   1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
          :::||||||||||||            ::::::|||||||||   |||:::||||||::::::
Db    494 GTTCAGTGTCCTAAAGGACCTGACACCCTATGGCAGTGCCCATCATCTCCATGGAAGAAG 553

Qy   1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
          |||::::::||||||:::||||||||||||||||||||   ::::::||||||::::::
Db    554 AGACTGGCCAGCCCCTCAGAGGAGACATGGATCACATGTGCCAACAAAATAAGACTTCAA 613

Qy   1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
             |||:::|||   ||||||   ||||||||||||||||||:::   |||||||||||||||
Db    614 GAAGGNAACACTAATTGTTCTNGACGTGTGGAGATCTGGTACGGNGGTTCCTGGGGCACT 673

Qy   1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCys 1284
          ||||||||||||||||||||||||      |||::::||||||||||
Db    674 GTGTGTGACGACTCCTGGGACTTGAA-AATGCNCAAGTGGTGTGC 717
```

```
RESULT 8
AL529438/c
LOCUS       AL529438                554 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL529438 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD006YO05 3
            prime, mRNA sequence.
ACCESSION   AL529438
VERSION     AL529438.1  GI:12792931
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DD006YO05"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax.: (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      132 a    166 c    127 g    110 t     19 others
ORIGIN

Alignment Scores:
Pred. No.:               4.57e-80      Length:          554
Score:                   895.00        Matches:         164
Percent Similarity:      90.81%        Conservative:    4
Best Local Similarity:   88.65%        Mismatches:      16
Query Match:             11.00%        Indels:          1
DB:                      9             Gaps:            0

US-09-759-130B-381 (1-1453) x AL529438 (1-554)

Qy   1189 AsnGlyValValSerLeuAlaProLeu-SerLysThrGlySerGlyPheMetTrpValAs 1208
          |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    554 AATGGAGTTATCAGCCTCGCCCCTTTATTCTAAGACAGGSTCTGGTTTCATGTGGGTGGA 495

Qy   1208 pAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGl 1228
          |   ||||||||||||   ||||||:::||||||||||||||||||||||||||||||||
Db    494 TRACATTCARTGTCCTRAAACGCATGTCTCCATATGGCAGTGCCTSTCTGCCCCATGGGA 435

Qy   1228 uArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgVa 1248
          ||||||   ||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    434 GCGAAGASTCTCCAGCCCVGCAGAAGAGACCTGGCTMACATGTGAAGATAGAATAAGAGT 375

Qy   1248 lArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGl 1268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 GCGTGGAGGAGACACCGAGTGCTCTGGGAGAGTGGAGATCTGGCACGCAGGCTCCTGGGG 315

Qy   1268 yThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGl 1288
          ||||||||||||   ||||||||||||||||||   ||||||   ||||||||||||
Db    314 CACAGTGTGTGATVACTCCTGGGACCTGGCCGAGGCGGNAGTGGTGTNTCAGCAGCTGGG 255

Qy   1288 yCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIl 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 CTGTGGCTCTGCTCTNGCTGCCCTGAGGGACGCTTCGTTTGGCCAGGGGACTGGAACCNT 195

Qy   1308 eTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLy 1328
           ||||||   |||||||||||||||   ||||||||||||||||||||||||||||||
Db    194 CTGGTTGGNTTACATGCGGTGCAAAGGAANTNAGTCATTTCTATGGGACTGTCACGCCAA 135

Qy   1328 sProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGl 1348
          ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db    134 ACCCTGGGGACAGAGTGACTGTGGACACAAGGGAGANGCTGGCGTTAGGTGCTCTGGACA 75

Qy   1348 nSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePh 1368
          |||||||   |||||||||||||||||||||||||||   ||||||||||||||:::||
Db     74 GTCGCTGNAATCACTGAATGCCTCCTCAGGTCATTTAGGACTTATTTTATCCAGTGTCTT 15

Qy   1368 eGlyLeuLeuLeu 1372
```

```
                 |||||||    |||
Db        14 TGGGCTCCKTCTC 2

RESULT 9
BI838644
LOCUS       BI838644                 930 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603086185F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5225205 5',
            mRNA sequence.
ACCESSION   BI838644
VERSION     BI838644.1  GI:15950194
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11566  row: j  column: 22
            High quality sequence stop: 731.
FEATURES             Location/Qualifiers
     source          1..930
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5225205"
                     /clone_lib="NIH_MGC_120"
                     /lab_host="DH10B"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT      236 a    207 c    274 g    213 t
ORIGIN

Alignment Scores:
Pred. No.:              1.07e-73        Length:         930
Score:                  836.50          Matches:        163
Percent Similarity:     66.88%          Conservative:   47
Best Local Similarity:  51.91%          Mismatches:     98
Query Match:            10.28%          Indels:         7
DB:                     13              Gaps:           3

US-09-759-130B-381 (1-1453) x BI838644 (1-930)

Qy       940 GlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSer 959
                 ||||||    ||||||    ||||||||||||   ||||||          |||:::    :::
Db         2 CAGCTTAAATGTGGAGTTGCCCTTTCTACCCCAGGAGGAGCACGTTTGGAAAAGGAAAT 61

Qy       960 ValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCys 979
                :::::::|||   |||   |||||||||   |||   |||          :::   :::|||
Db        62 GGTCAGATCTGGAGGCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTGT 121

Qy       980 GlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCys 999
                   :::|||    |||||||||        |||                ||||||||||||
Db       122 CCTGTAACTGCTCTAGGTGCTTCATTATGTCCTTCAGAGCAAGTGGCCTCTGTAATCTGC 181

Qy      1000 ThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeu 1019
             :::|||:::    :::|||   |||         |||    ::::::           |||

Db       182 TCAGGAAACCAGTCCCAAACACTGTCCTCGTGCAATTCATCGTCTTTGGGCCCAACAAGG 241

Qy      1020 SerAlaValProGluGlySer-AlaLeuIleCysLeuGluAspLysArgLeuArgLeuVa 1039
                   :::||||||    :::  |||:::    |||:::|||        :::|||||||||||
Db       242 CCTACCATTCCAGAAGAAACGTGCTGTGGCCTGCATAGAGAGTGGTCAACTTCGCCTGGT 301

Qy      1039 lAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyTh 1059
             |:::|||        |||||||||||||||||||||||||||||:::|||     ||||||||
Db       302 AAATGGAGGAGGTCGCTGTGCTGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCAC 361

Qy      1059 rIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCy 1079
             |||||||||||||   |||||||||||||||||||||||||||||:::::::||||||||
Db       362 CATCTGTGATGACAGCTGGGACCTGAGTGATGCCCACGTGGTTTGCAGACAGCTGGGCTG 421

Qy      1079 sGly-ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleT 1099
             ||||     |||    |||||·|||    |||||||||||||||||||||::·::|||||||||
Db       422 TGGACGAGGCCATTAATGCCACTGGTTCTGCTCATTTCGGGGAAGGAACAGGGCCCATCT 481

Qy      1099 rpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGln-CysProSerArg 1118
             |||||·|||::::::    |||      |||     ||||||      :::|||||| |||    |||
Db       482 GGCTGGATGAGATGAAATGCAATGGAAAAGAATCCCGCATTTGGCAGTTGCCATTCACAC 541

Qy      1119 GlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPhe 1138
             ||||||||||||      :::||||||||||||||||||||||||||||||||||||||||||
Db       542 GGCTGGGGGCAGCAAAATTGCAGGCACAAGGAGGATGCGGGAGTTATCTGCTCAGAATTC 601

Qy      1139 ThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluVal 1158
                :::||||||||||    ||||||           |||:::|||||||||||||||||||||||
Db       602 ATGTCTCTGAGACTGACCAGTGAAGCCAGCAGAGAGGCCTGTGCAGGCCGTCTGGAAGTT 661

Qy      1159 PheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGly 1178
             ||||||||||||    ||||||:::|||||||:::      :::::::::                |||
Db       662 TTTACAATGGAGCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAACCA---TGTGGT 718

Qy      1179 IleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSer 1198
             :::|||||||||||||||    |||    :::|||         ::::::|||          |||
Db       719 GTGGTGTGCAGGCAGCTGAGCTGTGCAGACAACAG-GAAATCAACCTGATT---CTTAGA 774

Qy      1199 LysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSer 1218
             :::                  |||||||||::::::::    |||    ||·|                |||
Db       775 CAAGTGCATGTCCATTCCATGTGGGTGGAAAATGTCACGTGTGCCAAAGGACCTGACAGC 834

Qy      1219 IleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThr 1238
                   |||          :::||||||||||:::    :::      :::|||
Db       835 TGTGGACAGTTGCCATTATCTCCATGGGAGAAGACACTTGGCAACCCCGGA---GGAAAC 891

Qy      1239 TrpIleThrCysGluAspArgIleArgValArgGly 1250
                :::||||||:::                   |||
Db       892 TGACTCACATGTGACACAGAATACAATCAAGAAGGC 927

RESULT 10
AL540904/c
LOCUS       AL540904                1153 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL540904 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE004YI19 3 prime
            , mRNA sequence.
ACCESSION   AL540904
VERSION     AL540904.1  GI:12871461
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1153)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..1153
                     /organism="Homo sapiens"
```

```
                    /db_xref="taxon:9606"
                    /clone="CS0DE004YI19"
                    /clone_lib="LTI_FL002_PL1"
                    /lab_host="DH10B"
                    /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
                    cDNA was primed with a NotI-oligo(dT) primer. Five prime
                    end enriched, double-stranded cDNA was digested with Not I
                    and cloned into the Not I and Eco RV sites of the
                    pCMVSPORT 6 vector.  Library was constructed by Life
                    Technologies. Contact : Feng Liang Life Technologies, a
                    division of Invitrogen 9800 Medical Center Drive Rockville
                    , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                    fliang@lifetech.com URL :
                    http://fulllength.invitrogen.com"
BASE COUNT      292 a    300 c    228 g    313 t     20 others
ORIGIN

Alignment Scores:
Pred. No.:              1.83e-66       Length:        1153
Score:                  767.00         Matches:       146
Percent Similarity:     65.45%         Conservative:  51
Best Local Similarity:  48.50%         Mismatches:    88
Query Match:            9.42%          Indels:        16
DB:                     9              Gaps:          5

US-09-759-130B-381 (1-1453) x AL540904 (1-1153)

Qy   1164 TrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGln 1183
          ||||||:::||||||:::|||:::::::::              |||:::|||||||||||||
Db   1153 TGGGGCACTGTTGGCAAGAGAAGCATGTCTGAAMCCACTGTGGGTGTGGTGTGCAGGCAG 1094

Qy   1184 LeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySerGly 1203
          ||||||||   :::   |||   :::        |||   |||   |||         |||
Db   1093 CTGGGCTGTGCAGACAAAGGGAAAATCAMCCCTGCATCTTTAGACAAGGCCATGTCCATT 1034

Qy   1204 PheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeu 1223
             |||||||||||:::::||||||||||||              :::    |||||||||
Db   1033 CCCATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGBTGTGGCAGTGCCCA 974

Qy   1224 SerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGlu 1243
          |||:::||||||||:::|||:::|||::::::||||||:::||||||||||||||||||:::
Db    973 TCATCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGAC 914

Qy   1244 AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHis 1263
          ::::::||||||::::::   |||   |||   ||||||||||||||||||||||||||||
Db    913 AACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCAT 854

Qy   1264 AlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValVal 1283
             ||||||||||||||||||||||||||||||||||||||||   :::|||:::||||||
Db    853 GGAGGTTCCTGGGGGACAGTGTGTGATGATTCTTGGGACTTGGACGATGCTCAGGTGGTG 794

Qy   1284 CysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGln 1303
          ||||||||||||||||||||   |||         |||   :::::::|||   |||||||||
Db    793 TGTCAACAACTTGGCTGTGGTCCAGCTTTKAAAGCATTCAAAGAAGCAGAGTTTGGTCAG 734

Qy   1304 GlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrp 1323
          ||||||||   |||||||||::::::::::::|||||||||||||||||   ||||||
Db    733 GGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGG 674

Qy   1324 AspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyVal 1343
          ||||||   |||:::   ||||||   |||:::||||||||||||||||||||||   |||
Db    673 GATTGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTG 614

Qy   1344 ArgCysSerGlyGlnSerLeuLys------------------SerLeuAsnAlaSerSer 1357
             |||:::   |||::::::           :::           ||||||
Db    613 AATTGCACAGATATTTCAGTGCAGAAAACCCCACAAAAAGCCACAACAGGTCGCTCATCC 554

Qy   1358 GlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIle 1377
                :::                 |||   |||::::::||||||   :::|||:::
Db    553 CGTCAGTCATCCTTTATKGCAGTCGGGATCCTTGGGGTTGTTCTGTTGGCCATTTTCGTC 494

Qy   1378 ------LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgVal 1395
                ||||||||      |||   |||:::|||             |||   |||
Db    493 GCATTATTCTTCTTGACTAAAAAGCGAAGACAGAGACAGCMS---------CTTGCAGTT 443

Qy   1396 SerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeu 1415
             |||             :::   :::      :::   ||||||   :::||||||
Db    442 CCCCCAAGAGGAGAGAACTTAGTCCACCAAATTCAATACCGGGAGATGAANTCTTGCCTG 383

Qy   1416 LysArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGlu 1435
             :::|||               :::|||::::::         :::|||
Db    382 AAKGCAGATGATCTGGACCTAATGAATTCCTCAGAAAAT------TCCCATGAGTCAGCT 329

Qy   1436 AspAlaSerAspThrSerLeuLeuGlyVal---------LeuProAlaSerGluAlaThr 1452
          |||   |||.       |||:::   |||            ||||||   |||
Db    328 GATTTCAGTGCTGCTGAACTAATTTCTGTGTCTAAATTTCTTCCTATTTCTGGAATGGAA 269

Qy   1453 Lys 1453
          |||
Db    268 AAG 266

RESULT 11
BI836459
LOCUS       BI836459                723 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603082833F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5221905 5',
            mRNA sequence.
ACCESSION   BI836459
VERSION     BI836459.1  GI:15948009
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11558  row: a  column: 10
            High quality sequence stop: 708.
FEATURES             Location/Qualifiers
     source          1..723
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5221905"
                     /clone_lib="NIH_MGC_120"
                     /lab_host="DH10B"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT      173 a    160 c    237 g    153 t
ORIGIN

Alignment Scores:
Pred. No.:              5.24e-64       Length:        723
Score:                  739.00         Matches:       144
Percent Similarity:     73.88%         Conservative:  37
Best Local Similarity:  58.78%         Mismatches:    54
Query Match:            9.08%          Indels:        10
DB:                     13             Gaps:          0

US-09-759-130B-381 (1-1453) x BI836459 (1-723)
```

```
Qy    1010 CysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIle 1029
           |||   ::::::         |||            :::||||||   |||||||:::
Db       2 TGCAATTCATCGTCTTTGGGCCCAACAAGGCCTACCATTCCAGAAGAAAGTGCTGTGGCC 61

Qy    1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049
           |||:::|||         :::|||||||||||||:::|||      |||||||||||||||||
Db      62 TGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGAGGAGGTCGCTGTGCTGGGAGAGTA 121

Qy    1050 GluIleTyrHisAspGlyPheTrpGlyThrIle-CysAspAspGlyTrpAspLeuSerAs 1069
           ||||||||||||:::|||   ||||||||||||  |||||||||   ||||||||||||||
Db     122 GAGATCTATCATGAGGGCTCCTGGGGCACCATCGTGTGATGACAGCTGGGACCTGAGTGA 181

Qy    1069 pAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThr-ValSerA 1089
           ||||||||||||||||:::::::||||||||||||   |||    |||||||||     ||||
Db     182 TGCCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACGTGGTTCTG 241

Qy    1089 laHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThr- 1108
           |||||   ||||||||:::||||||||||||||||||||:::::::   |||    ||||||
Db     242 CTCATTTAGGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAAATGCAATGGAACAA 301

Qy    1109 -GluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLy 1128
            :::|||    :::|||||||||    |||    |||||||||||    :::||||||||||||
Db     302 GAAGTCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGGCAGCAAAATTGCAGGCACAA 361

Qy    1128 sGluAspAlaGlyValIleCysSerGlu-PheThrAlaLeuArgLeuTyrSerGluThrG 1148
           ||||||||||||||||||||||||||| |||     :::|||||||||    ||||||
Db     362 GGAGGATGCGGGAGTTATCTGCTCAGAAGTTCATGTCTCTGAGACTGACCAGTGAAGCCA 421

Qy    1148 luThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValG 1168
                 |||:::||||||||||||||||||||||||||||||||||||   ||||||:::||||
Db     422 GCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTTACAATGGAGCTTGGGGCACTGTTG 481

Qy    1168 lyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyG 1188
           ||:::     :::::::::        |||:::|||||||||||||||||||||||   :
Db     482 GCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCAG 541

Qy    1188 luAsnGlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValA 1208
           ::  .|||    ::::::   |||    |||    |||         |||        ||||||||||
Db     542 ACGAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTGG 601

Qy    1208 sp-AspIleGlnCysProLysThr-HisIleSerIleTrpGlnCys--LeuSerAlaPro 1226
           ||      :::|||||||||            ::::::||||||||||      |||:::|||
Db     602 ACAAGTGTTCAGTGTCCAAGCAGGACCTGACACGCTGTGGCAGTGCCCGAGTCATCTCCA 661

Qy    1227 Trp-GluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAsp 1244
           |||  ::::::|||::::::      |||:::||||||||||||||||||||:::|||
Db     662 TGGGCAGAAGAGACTGGCCAGGCCCTCGGAGGAGACCTGGATCACATGTGACGAC 716
```

RESULT 12
BI009838
LOCUS   BI009838   644 bp   mRNA   linear   EST 13-JUN-2001
DEFINITION MR4-RT0048-170401-008-b11 RT0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI009838
VERSION   BI009838.1 GI:14413909
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 644)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   MEDLINE 20202663
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-RT0048-170401-008-b11&t3=2001-04-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 33
            High quality sequence stop: 538.
FEATURES   Location/Qualifiers
     source   1..644
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="RT0048"
                     /dev_stage="Adult"
                     /note="Organ: kidney_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT   156 a   128 c   210 g   150 t
ORIGIN

```
Alignment Scores:
Pred. No.:               5.69e-62      Length:          644
Score:                   718.00        Matches:         121
Percent Similarity:      79.90%        Conservative:    34
Best Local Similarity:   62.37%        Mismatches:      39
Query Match:             8.82%         Indels:          1
DB:                      13            Gaps:            0
```

US-09-759-130B-381 (1-1453) x BI009838 (1-644)

```
Qy    1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169
           |||:::|||||||||   ||||||||||||||||||||   ||||||:::||||||:::
Db      44 GAGGCCTGTGCAGGGGT-CTGGAAGTTTTTTACAATGGAGCTTGGGGCACTGTTGGCAAG 102

Qy    1170 ArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
              :::::::::               |||:::|||||||||||||||||||||   :::
Db     103 AGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCAGACAAA 162

Qy    1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
           |||    ::::::   |||    |||    |||         |||         |:||||||||||:::
Db     163 GGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTGGACAAT 222

Qy    1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
           :::||||||||||||           ::::::||||||||||   |||:::||||||||||:::
Db     223 GTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGGGAGAAG 282

Qy    1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
           |||::::::||||||:::|||||||||||||||||||||||:::::::::||||||::::::
Db     283 AGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAACAAGATAAGACTTCAG 342

Qy    1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
              |||   |||   ||||||||||||||||||||||||||   ||||||||||||||
Db     343 GAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGAGGTTCCTGGGGGACA 402

Qy    1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
           ||||||||||||||||||||||||   :::|||:::|||||||||||||||||||||||||||
Db     403 GTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTTGGTTGT 462

Qy    1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
           |||   ||||||   |||     ::::::|||   ||||||||||||||||||   ||||||
Db     463 GGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCGATATGG 522

Qy    1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
           |||::::::::::::|||   ||||||||||||   ||||||||||||         :::
```

```
Db     523 CTCAATGAAGTGAAGTGCACAGGGAATGAGTCTTCCTTGTGGGATTGTCCTGGCAGACGC 582

Qy    1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyVal 1343
           ||||||   |||:::|||||||||   |||||||||   |||
Db     583 TGGGGGCATAGTGAGTGTGGGCACACGGAAGACGCTGCAGTG 624
```

RESULT 13
BQ327934

```
LOCUS       BQ327934                 585 bp    mRNA    linear   EST 17-MAY-2002
DEFINITION  MR4-RT0048-070501-010-g04 RT0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ327934
VERSION     BQ327934.1  GI:20945766
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-RT0048-
            070501-010-g04&t3=2001-05-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 24
            High quality sequence stop: 584.
FEATURES             Location/Qualifiers
     source          1..585
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="RT0048"
                     /dev_stage="Adult"
                     /note="Organ: kidney_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      144 a    122 c    186 g    133 t
ORIGIN

Alignment Scores:
Pred. No.:               2.52e-60      Length:          585
Score:                   701.00        Matches:         122
Percent Similarity:      80.83%        Conservative:    34
Best Local Similarity:   63.21%        Mismatches:      37
Query Match:             8.61%         Indels:          1
DB:                      14            Gaps:            0

US-09-759-130B-381 (1-1453) x BQ327934 (1-585)

Qy    1128 LysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThr 1147
           :::||||||||||||||   ||||||||||||   :::|||||||||   ||||||
Db       3 AGGGAGGATGCGGGAGTTAC-TGCTCAGAATTCATGTCTCTGAGACTGACCAGTGAAGCC 61

Qy    1148 GluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerVal 1167
                 |||:::|||||||||||||||||||||||||||||||||   ||||||:::|||
Db      62 AGCAGAGAGGCCTGTGCAGGGCGTCTGGAAGTTTTTACAATGGAGCTTGGGGCACTGTT 121

Qy    1168 GlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGly 1187
           |||:::   :::::::::            |||:::||||||||||||||||||||||
Db     122 GGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGGCTGTGCA 181

Qy    1188 GluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpVal 1207
           :::   |||   ::::::   |||   |||   |||      |||      ||||||||
Db     182 GACAAAGGGAAAATCAACCCTGCATCTTTAGACAAGGCCATGTCCATTCCCATGTGGGTG 241

Qy    1208 AspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp 1227
           |||::::::|||||||||||         ::::::|||||||||   |||:::||||||
Db     242 GACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGG 301

Qy    1228 GluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArg 1247
           |||:::|||::::::||||||:::||||||||||||||||||||||::::::::::||||||
Db     302 GAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAACAAGATAAGA 361

Qy    1248 ValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrp 1267
           ::::::   |||   |||   ||||||||||||||||||||||||||||   |||||||||
Db     362 CTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGAGGTTCCTGG 421

Qy    1268 GlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeu 1287
           ||||||||||||||||||||||||||||||   :::|||:::||||||||||||||||||
Db     422 GGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTT 481

Qy    1288 GlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThr 1307
           |||||||||   ||||||   |||   :::::::|||   ||||||||||||||||||
Db     482 GGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCG 541

Qy    1308 IleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
           |||||||||:::::::::::::|||   ||||||||||||
Db     542 ATATGGCTCAATGAAGTGAAGTGCACAGGGAATGAGTCT 580
```

RESULT 14
BE502724

```
LOCUS       BE502724                 653 bp    mRNA    linear   EST 04-AUG-2000
DEFINITION  hz78d12.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3214103 3'
            similar to TR:Q07898 Q07898 M130 ANTIGEN PRECURSOR. ;, mRNA
            sequence.
ACCESSION   BE502724
VERSION     BE502724.1  GI:9705132
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 653)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 483.
FEATURES             Location/Qualifiers
     source          1..653
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3214103"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
```

```
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      171 a    134 c    191 g    155 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.26e-59      Length:        653
Score:                  695.00        Matches:       124
Percent Similarity:     70.51%        Conservative:  29
Best Local Similarity:  57.14%        Mismatches:    64
Query Match:            8.54%         Indels:        0
DB:                     10            Gaps:          0

US-09-759-130B-381 (1-1453) x BE502724 (1-653)

Qy    684 AspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysVal 703
Db      3 GATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGCAGCCGCTGTGCTGGGACAGTT 62

Qy    704 GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle 723
Db     63 GAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGGACTGAAAGAA 122

Qy    724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743
Db    123 GCTGATGTGGTTTGCAGGCAGCTGGGATGTGGATCTGCACTCAAAACATCTTATCAAGTG 182

Qy    744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763
Db    183 TACTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTAACGGAAATGAA 242

Qy    764 AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783
Db    243 ACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAA 302

Qy    784 GluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMetPro 803
Db    303 GAAGCCAAAATTACCTGCTCAGCCCACAGGGAACCCAGACTGGTTGGAGGGGACATTCCC 362

Qy    804 CysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAsp 823
Db    363 TGTTCTGGACGTGTTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGATTCGGAC 422

Qy    824 PheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSer 843
Db    423 TTCTCTCTGGAAGCTGCCAGCGTTCTATGCAGGGAATTACAGTGTGGCACAGTTGTCTCT 482

Qy    844 LeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGln 863
Db    483 ATCCTGNGGGGAGCTCACTTTGGAGAAGGAAATGGACAGATCTGGGCTGAAGAATTNCAT 542

Qy    864 CysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThr 883
Db    543 TGTGAGGGACATGAGTCCCATCTTTCACTCTGCCCAATAGCACCGCCCCCAGAAAGGACT 602

Qy    884 CysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
Db    603 TGTATCCTCAGCAAGGATGTTGGAGTAGTCTGCTTAAGATCACAGAACATC 653

RESULT 15
BI009795/c
LOCUS       BI009795                 616 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  MR4-RT0048-170401-007-b01 RT0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI009795
VERSION     BI009795.1  GI:14413866
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-RT0048-
            170401-007-b01&t3=2001-04-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1..616
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="RT0048"
                     /dev_stage="Adult"
                     /note="Organ: kidney_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      144 a    193 c    129 g    150 t
ORIGIN

Alignment Scores:
Pred. No.:              4.76e-58      Length:        616
Score:                  679.00        Matches:       112
Percent Similarity:     79.89%        Conservative:  35
Best Local Similarity:  60.87%        Mismatches:    37
Query Match:            8.34%         Indels:        0
DB:                     13            Gaps:          0

US-09-759-130B-381 (1-1453) x BI009795 (1-616)

Qy   1165 GlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeu 1184
Db    588 GGCACTGTTGCCAAGAGTAGCATGTCTGAAACCACTGTGGGTTTGGTTTGCAGGCAGCTG 529

Qy   1185 GlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPhe 1204
Db    528 GGCTGTGCAGACAAAGGGAAAATCAACCCCGCATCTTTAGACAAGGCCATGTCCATTCCC 469

Qy   1205 MetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSer 1224
Db    468 ATGTGGGTGGACAATGTTCAGTGTCCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCA 409

Qy   1225 AlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAsp 1244
Db    408 TCTCCATGGGAGAAGAGACTGGCCAGCCCCTCGGAGGAGACCTGGATCACATGTGACAAC 349
```

```
Qy   1245 ArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAla 1264
          :::||||||:::::   |||   |||   |||||||||||||||||||||||||
Db    348 AAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGAGATCTGGCATGGA 289

Qy   1265 GlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCys 1284
          ||||||||||||||||||||||||||||||||||||||||||   :::|||:::|||||||||
Db    288 GGTTCCTGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGT 229

Qy   1285 GlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGly 1304
          |||||||||||||||||   ||||||   |||   ::::::|||   |||||||||||
Db    228 CAACAACTTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGG 169

Qy   1305 ThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAsp 1324
          ||||||   ||||||||:::::::::::||||||||||||||||||   |||||||||
Db    168 ACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAAGGGAATGAGTCTTCCTTGTGGGAT 109

Qy   1325 CysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArg 1344
          |||   |||:::   ||||||   |||:::||||||||||||||||||||   |||
Db    108 TGTCCTGCCAGACGCTGGGGCCATAGTGAGTGTGGGCACAAGGAAGACGCTGCAGTGAAT 49

Qy   1345 CysSerGlyGln 1348
             :::|||:::
Db     48 CGCACAGGTAAG 37
```

Search completed: May 12, 2003, 17:31:28
Job time : 2883 secs

OM protein - nucleic search, using frame_plus_p2n model

```
Run on:          May 12, 2003, 10:35:25 ; Search time 79 Seconds
                                          (without alignments)
                                          5640.523 Million cell updates/sec

Title:           US-09-759-130B-381
Perfect score:   8138
Sequence:        1 MMLPQNSWHIDFGRCCCHQN..........CEDASDTSLLGVLPASEATK 1453

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09759130/runat_06052003_123555_23377/app_query.fasta_1.1607
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759130_@CGN_1_1_41_@runat_06052003_123555_23377 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2435	29.9	5802	4	US-09-341-587-4	Sequence 4, Appli
	2	1142.5	14.0	4360	1	US-08-470-350B-1	Sequence 1, Appli
	3	785.5	9.7	29598	4	US-09-341-587-6	Sequence 6, Appli
	4	715.5	8.8	2001	4	US-09-341-587-2	Sequence 2, Appli
	5	691.5	8.5	2178	3	US-09-034-916-1	Sequence 1, Appli
	6	634.5	7.8	2262	3	US-09-276-400-3	Sequence 3, Appli
	7	634.5	7.8	2262	4	US-09-448-076-3	Sequence 3, Appli
	8	634.5	7.8	2262	4	US-09-702-572-3	Sequence 3, Appli
	9	634.5	7.8	2920	3	US-09-276-400-1	Sequence 1, Appli
	10	634.5	7.8	2920	4	US-09-448-076-1	Sequence 1, Appli
	11	634.5	7.8	2920	4	US-09-702-572-1	Sequence 1, Appli
	12	618.5	7.6	28720	4	US-09-341-587-7	Sequence 7, Appli
	13	314	3.9	2285	1	US-08-477-674-9	Sequence 9, Appli
	14	314	3.9	2285	1	US-08-473-791-9	Sequence 9, Appli
	15	314	3.9	2285	2	US-08-316-714-9	Sequence 9, Appli
	16	314	3.9	2285	3	US-08-473-673-9	Sequence 9, Appli
	17	290.5	3.6	1868	1	US-08-392-367B-1	Sequence 1, Appli
	18	290.5	3.6	1868	3	US-08-893-467A-1	Sequence 1, Appli
	19	290	3.6	2028	1	US-08-453-117-1	Sequence 1, Appli
	20	290	3.6	2028	2	US-08-948-222-1	Sequence 1, Appli
	21	290	3.6	2028	2	US-08-973-145-1	Sequence 1, Appli
	22	290	3.6	2028	5	PCT-US96-08081-1	Sequence 1, Appli
	23	290	3.6	2037	1	US-08-154-365-1	Sequence 1, Appli
	24	288.5	3.5	1560	2	US-08-794-795-5	Sequence 5, Appli
	25	288.5	3.5	1560	4	US-09-249-200-5	Sequence 5, Appli
	26	284.5	3.5	1703	2	US-08-794-795-1	Sequence 1, Appli
	27	284.5	3.5	1703	4	US-09-249-200-1	Sequence 1, Appli
c	28	282	3.5	410	4	US-09-221-298-83	Sequence 83, Appl
	29	268	3.3	1588	6	5510466-3	Patent No. 5510466
	30	248.5	3.1	9723	1	US-08-083-590A-21	Sequence 21, Appl
	31	248.5	3.1	9723	3	US-08-532-384-21	Sequence 21, Appl
	32	244	3.0	3955	4	US-09-214-278-4	Sequence 4, Appli
	33	243.5	3.0	4315	3	US-08-882-046-3	Sequence 3, Appli
	34	243.5	3.0	8257	4	US-09-484-970B-65	Sequence 65, Appl
	35	239.5	2.9	4483	3	US-08-611-729A-7	Sequence 7, Appli
	36	239	2.9	4464	2	US-08-400-159-7	Sequence 7, Appli
	37	208.5	2.6	5804	4	US-09-369-364A-12	Sequence 12, Appl
	38	207	2.5	7286	5	PCT-US95-11684-1	Sequence 1, Appli
	39	200	2.5	6677	4	US-08-939-366-27	Sequence 27, Appl
	40	200	2.5	6677	4	US-09-467-997-6	Sequence 6, Appli
c	41	196.5	2.4	4394	2	US-08-750-152A-1	Sequence 1, Appli
	42	193.5	2.4	16382	4	US-08-718-388-8	Sequence 8, Appli
	43	186	2.3	4208	4	US-09-214-278-6	Sequence 6, Appli
	44	185	2.3	5499	3	US-08-479-722B-1	Sequence 1, Appli
	45	185	2.3	5502	5	PCT-US95-02251-17	Sequence 17, Appl

ALIGNMENTS

```
RESULT 1
US-09-341-587-4
; Sequence 4, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
;  APPLICANT: Mollenhauer, Jan
;  TITLE OF INVENTION: Protein Containing an SRCR Domain
;  FILE REFERENCE: 4121-108
;  CURRENT APPLICATION NUMBER: US/09/341,587
;  CURRENT FILING DATE: 1999-08-31
;  EARLIER APPLICATION NUMBER: PCT/DE98/00096
;  EARLIER FILING DATE: 1998-01-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 5802
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-4

Alignment Scores:
Pred. No.:                1.93e-215      Length:         5802
Score:                    2435.00        Matches:        533
Percent Similarity:       46.58%         Conservative:   155
Best Local Similarity:    36.09%         Mismatches:     415
Query Match:              29.92%         Indels:         374
DB:                       4              Gaps:           28

US-09-759-130B-381 (1-1453) x US-09-341-587-4 (1-5802)

Qy      43 GlyThrAsp------LeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
           |||:::|||      |||   ||||||||||||||||||||||||    |||    |||
Db     389 GGATCTGATTCTGGTTTGGCCCTGAGGCTGGTGAATGGAGATGGCAGGTGTCAGGGCCGA 448

Qy      61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
           ||||||:::    :::::::|||   ||||||||||||||||||||    |||:::|||
```

```
Db    449 GTGGAGATCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGACAGCTGGGACACCAAT 508

Qy     81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMet--------- 97
Db    509 GATGCCAACGTGGTCTGTAGGCAGCTGGGTTGTGGCTGGGCCATGTCAGCTCCAGGAAAT 568

Qy     98 PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCys 117
Db    569 GCCTGGTTTGGCCAGGGC---------TCAGGACCCATTGCCCTGGATGATGTGCGCTGC 619

Qy    118 TyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCys 137
Db    620 TCAGGACACGAATCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAACTGT 679

Qy    138 TyrHisGlyGluAspValGlyValAsnCysTyr--------------------------- 148
Db    680 GGCCATGGTGAAGATGCTGGTGTTATCTGCTCAGCTGCCCAGCCTCAGTCAACACTCAGG 739

Qy    149 ------------------------------------------------GlyGluAlaAsn 152
Db    740 CCAGAAAGTTGGCCTGTCAGGATATCACCACCTGTACCCACAGAAGGATCTGAATCCAGT 799

Qy    153 LeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPhe 172
Db    800 TTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGGTCCTATAC 859

Qy    173 GlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValVal 192
Db    860 CGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGGTC 919

Qy    193 CysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla 212
Db    920 TGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAG 979

Qy    213 ValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrp 232
Db    980 GGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCCTACCTGTGG 1039

Qy    233 AsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeu 252
Db   1040 AGCTGCCCCCACAATGGCTGGCTCACCCACAACTGTGGCCATAGTGAAGACGCTGGTGTC 1099

Qy    253 ThrCysTyr--------------------------------------------------- 255
Db   1100 ATCTGCTCAGCTCCCCAGTCCCGGCCGACACCCAGCCCAGATACTTGGCCGACCTCACAT 1159

Qy    256 ---------------AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArg 270
Db   1160 GCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGG 1219

Qy    271 CysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLys 290
Db   1220 TGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGTGATGATAGC 1279

Qy    291 TrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHis 310
Db   1280 TGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGCTGTGGCTGGGCCACGTCA 1339

Qy    311 PheAlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSer 330
Db   1340 GCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTGGATGACGTGCGC 1399

Qy    331 CysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAsp 350
Db   1400 TGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCATAAC 1459

Qy    351 CysLeuHisGlnAsnAspValSerValIleCysSerAspGlyAlaAspLeuGluLeuArg 370
Db   1460 TGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGCTGCC------------------ 1501

Qy    371 LeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrp 390
Db   1502 ---------------------------------------------------CACTCCTGG 1510

Qy    391 TrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu 410
Db   1511 TCGACG------------------------------------------------------ 1516

Qy    411 GlyCysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAsp 430
Db   1517 ---------------------------------------CCCAGTCCAGACACATTGCCG 1537

Qy    431 IleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyr 450
Db   1538 ACCATCACCTTGCCTGCATCGACAGTAGGATCTGAATCCAGT------------------ 1579

Qy    451 AspGlyLysAlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSerAsp 470
Db   1579 ------------------------------------------------------------ 1579

Qy    471 LysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGlu 490
Db   1580 ---------TTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAG 1630

Qy    491 ValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsnAla 510
Db   1631 GTCCTATACCAAGGCTCCTGGGGCACCGTGTGCGATGACAGCTGGGACACCAATGATGCC 1690

Qy    511 AlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThrTyr 530
Db   1691 AATGTCGTCTGCAGGCAACCGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCGG 1750

Qy    531 PheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSer 550
Db   1751 TTTGGTCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACACGAGTCT 1810

Qy    551 AsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAsp 570
Db   1811 TACCCGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATAGTGAAGAC 1870

Qy    571 ValIleValThrCysSerGlyAspAla-----------------------ThrTrp--- 581
Db   1871 GCTGGTGTCATCTGCTCAGCTTCCCAGTCCCGGCCAACACCTAGTCCAGACACTTGGCCA 1930

Qy    582 -------------------------------------GlyLeuArgLeuValGlyGly 588
Db   1931 ACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGA 1990

Qy    589 SerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrValCys 608
Db   1991 GGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGT 2050

Qy    609 AspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnLeuAspCysProSer 628
Db   2051 GATGACTACTGGGACACCAATGATGCCAATGTGGTTTGCAGGCAGCTGGGCTGTGGCTGG 2110

Qy    629 SerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAsp 648
Db   2111 GCCATGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATTGTCCTGGAT 2170

Qy    649 AspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGly 668
Db   2171 GATGTGCGCTGCTCAGGACATGAGTCCTATCTGTGGAGCTGCCCCCACAATGGCTGGCTC 2230

Qy    669 AsnAsnAspCysSerHisSerGluAspValGlyValIleCysSer--------------- 683
Db   2231 TCCCACAACTGTGGCCATCATGAAGACGCTGGTGTCATCTGCTCAGCTTCCCAGTCCCAG 2290

Qy    684 ---------------------------------------------------AspAlaSer 686
Db   2291 CCGACACCCAGCCCAGACACTTGGCCAACCTCACATGCATCAACAGCAGGATCTGAATCC 2350

Qy    687 AspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValGluValAsn 706
Db   2351 AGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTA 2410
```

```
Qy    707 ValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluVal 726
             :::|||:::   |||   :::|||   :::   |||    |||   |||   |||
Db   2411 TACCGAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTG 2470

Qy    727 ValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluProHisPheThr 746
          |||||||||||||||   ||||||   |||                  |||
Db   2471 GTTTGCAGGCAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGC 2530

Qy    747 GluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeu 766
          :::   :::         |||::::::   :::         |||:::|||   |||:::   |||
Db   2531 CAGGGTTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACATGAGTCCTATCTG 2590

Qy    767 TrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSer 786
          |||   |||               |||         |||      :::   :::|||
Db   2591 TGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCATGAAGACGCTGGT 2650

Qy    787 LeuIleCysSerAlaHisArg---GlnProArgLeuValGlyAlaAspMetProCysSer 805
          :::||||||||||||      :::   ||||||
Db   2651 GTCATCTGCTCAGCTTCCCAGTCCCAGCCG------------------------------ 2680

Qy    806 GlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSer 825
                                   |||||||||   :::
Db   2681 ------------ACACCCAGCCCAGACACTTGGCCAACC--------------------- 2707

Qy    826 LeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSer 845

Db   2707 ------------------------------------------------------------ 2707

Qy    846 ValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGlu 865
                                                            :::
Db   2708 ---------------------------------------------TCTCGTGCATCAACAGCA 2725

Qy    866 GlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIle 885
          |||||||||:::   |||||||||
Db   2726 GGATCTGAATCCACTTTGGCCCTG------------------------------------ 2749

Qy    886 HisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGly 905
                                                      |||||||||||||||
Db   2750 ---------------------------------------------AGACTGGTGAATGGA 2764

Qy    906 LysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCys 925
                :::|||   |||:::||||||:::         |||   ||||||::::::|||
Db   2765 GGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCAAGGCTCCTGGGGCACCGTGTGT 2824

Qy    926 AspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThr 945
          |||   :::||||||      ||||||   |||:::||||||||||||      ||||||
Db   2825 GATGACTACTGGGACACCAATGATGCCAACGTGGTCTGCAGGCAGCTGGGCTGTGGCTGG 2884

Qy    946 AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHis 965
          |||:::|||         |||         |||:::   |||         :::
Db   2885 GCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGAT 2944

Qy    966 ArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGly 985
                      |||   |||:::||||||   |||      :::|||
Db   2945 GATGTGCGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTC 3004

Qy    986 AlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGln 1005
          :::         |||   |||               |||||||||:::   :::   :::|||
Db   3005 TCCCACAACTGTGGCCATCATGAAGATGCTGGTGTCATCTGCTCAGCTGCTCAGTCCCAG 3064

Qy   1006 ------ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValPro 1023
                |||                        :::   |||   |||:::
Db   3065 TCAACGCCCAGGCCAGATACTTGGCTGACCACCAACTTACCGGCATTGACAGTAGGATCT 3124

Qy   1024 GluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSer 1043
          |||   |||                              ||||||||||||:::|||
Db   3125 GAATCCAGTTTG--------------------GCTCTGAGGCTGGTGAATGGAGGTGAC 3163

Qy   1044 ArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAsp 1063
          ||||||   ||||||||||||:::   :::   |||      |||||||||:::|||||||||
Db   3164 AGGTGTCGAGGCCGAGTGGAGGTCCTGTATCGAGGCTCCTGGGGAACCGTGTGTGATGAC 3223

Qy   1064 GlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPhe 1083
          ||||||   :::||||||:::||||||||||::::::|||||||||||    |||
Db   3224 AGCTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATG 3283

Qy   1084 AsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeu 1103
          :::|||      :::|||   ||||||:::||||||||||||||||   |||||||||:::
Db   3284 TCGGCCCCAGGAAATGCCCGGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGATGATGTG 3343

Qy   1104 AsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHis 1123
             |||:::|||   ||||||:::||||||   ||||||   :::||||||      |||
Db   3344 CGCTGCTCAGGGAATGAGTCCTACCTGTGGAGCTGCCCCCACAAAGGCTGGCTCACCCAC 3403

Qy   1124 AspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeu 1143
          :::|||   |||   ||||||||||||||||||||||||         |||   ::::::
Db   3404 AACTGTGGCCATCACGAAGACGCTGGTGTCATCTGCTCA---------GCCACCCAAATA 3454

Qy   1144 TyrSerGluThr------------------------------------GluThrGlu 1150
             |||   |||                                       :::
Db   3455 AATTCTACTACGACAGATTGGTGGCATCCAACAACTACAACCACTGCAAGACCCTCTTCA 3514

Qy   1151 SerCysAlaGlyArgLeuGluValPheTyr---AsnGlyThrTrpGlySer--------- 1166
          :::|||   |||   |||         ||||||   :::||||||:::   |||
Db   3515 AATTGTGGTGGCTTCTTA------TTCTATGCCAGTGGGACATTCTCCAGCCCATCCTAC 3568

Qy   1166 ------------------------------------------------------------ 1166

Db   3569 CCTGCATACTACCCCAACAATGCTAAGTGTGTTTGGGAAATAGAAGTGAATTCTGGTTAT 3628

Qy   1167 ---------ValGlyArgArgAsnIle--------------------------------- 1172
                   :::|||   |||:::
Db   3629 CGCATAAACCTGGGCTTCAGTAATCTGAAATTGGAGGCACACCATAACTGCAGTTTTGAT 3688

Qy   1173 ----------------------------ThrThrAlaIleAlaGlyIleValCys--- 1181
                                      ::::::   :::   |||   :::|||
Db   3689 TATGTTGAAATCTTTGATGGATCATTGAATAGCAGTCTCCTGCTGGGGAAAATCTGTAAT 3748

Qy   1182 ------ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu---------AlaPro 1196
                ||||||:::                  :::::::::            :::
Db   3749 GATACCAGGCAAATATTTACATCTTCTTACAACCGAATGACCATTCACTTTCGAAGTGAC 3808

Qy   1197 LeuSerLysThrGlySerGlyPheMet---TrpValAspAspIleGlnCysProLysThr 1215
          :::|||         :::|||||||||:::   |||   :::
Db   3809 ATCAGTTTCCAAAACACTGGCTTTTGGCTTGGTATAAC-------------------- 3847

Qy   1216 HisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAla 1235
                                                      |||   |||:::
Db   3848 -----------------------------------------------TCCTTCCCAAGC 3859

Qy   1236 GluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCys 1255
          :::   |||                        :::   :::                  |||
Db   3860 GATGCCACCTTG-----------AGGTTGGTCAATTTAAATCATCCTATGGTCTATGT 3907

Qy   1256 SerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrp 1275
          :::||||||||||||||||||:::|||   |||:::|||||||||||||||||||||||||||
Db   3908 GCCGGGCGTGTAGAAATTTACCATGGTGGCACCTGGGGGACAGTTTGTGATGACTCCTGG 3967

Qy   1276 AspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAla 1295
             :::   ||||||||||||||||||:::||||||||||||||||   |||::::::|||
Db   3968 ACCATTCAGGAAGCTGAGGTGGTCTGCAGACAGCTAGGGTGTGGACGTGCAGTTTCAGCC 4027

Qy   1296 LeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCys 1315
          |||   :::|||   ||||||   |||:::|||   |||   |||||||||:::   |||
Db   4028 CTTGGAAATGCATATTTTGGCTCTGGCTCTGGCCCCATCACCCTGGACGATGTAGAGTGC 4087

Qy   1316 LysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCys 1335
             |||   ||||||   ||||||   |||         :::   |||         :::|||
Db   4088 TCAGGGACGGAATCCACTCTCTGGCAGTGCCGGAACCGAGGCTGGTTCTCCCACAACTGT 4147

Qy   1336 GlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSer 1352
          |||:::|||||||||||||||   |||||||||         |||   :::
Db   4148 AATCATCGTGAAGATGCTGGTGTCATCTGCTCAGGAAACCATCTATCGACA 4198
```

RESULT 2
US-08-470-350B-1

```
; Sequence 1, Application US/08470350B
; Patent No. 5684126
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Xiao
;    APPLICANT:  Snyder, Solomon H
;    TITLE OF INVENTION:  Ebnerin: A Secreted von Ebner's Gland
;    TITLE OF INVENTION:  Protein Associated with Taste Buds
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patent In Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,350B
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wolffe, Susan A
;      REGISTRATION NUMBER:  33,568
;      REFERENCE/DOCKET NUMBER:  01107.48790
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4360 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus rattus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  94..3963
US-08-470-350B-1

Alignment Scores:
Pred. No.:              1.22e-95      Length:         4360
Score:                  1142.50       Matches:        331
Percent Similarity:     41.73%        Conservative:   136
Best Local Similarity:  29.58%        Mismatches:     343
Query Match:            14.04%        Indels:         312
DB:                     1             Gaps:           39

US-09-759-130B-381 (1-1453) x US-08-470-350B-1 (1-4360)

Qy    258 SerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeu 277
Db     10 TCTGGTTTGGCTGTGAGGCTGGTGAATGGAGGAGACAGGTGTCGGGGTCGCGTGGAGATC 69

Qy    278 LysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAsp 297
Db     70 CTTTACCAGGGTTCCTGGGGCACCATGTGTGATGACAGCTGGGACATCAATGATGCCAAC 129

Qy    298 ValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeu 317
Db    130 GTGGTGTGCAGGCAGCTGGGCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAGTTT 189

Qy    318 GlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPhe 337
Db    190 GGACAGGGTCTGGGTCCCATTGTTCTGGATGACGTGGCCTGTAGAGGACATGAGGCCTAT 249

Qy    338 LeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspVal 357
Db    250 CTGTGGAGCTGCTCCCACCGAGGCTGGCTCTCTCATAACTGTGGACATCAGGAGGATGCT 309

Qy    358 SerValIleCysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsn 377
Db    310 GGAGTGATCTGCTCAGAT------------------------TCTCAAACAAGCAGTCCC 345

Qy    378 CysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsn 397
Db    346 ACACCCGGT---------------------------TGGTGG---------AACCCCGGG 369

Qy    398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417
Db    370 GGCACAAATAACGATGTGATCTACGACACTCAAGAAACC--------------------- 408

Qy    418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437
Db    409 ---ACAGAAACTTCTCAAACAAGCAGTCCCACACCTGATTGGTGG--------------- 450

Qy    438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr 457
Db    451 ------AACCATGGGGGCACAATTAATGATGTGATCTATGAC---------ACTCAAGAA 495

Qy    458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477
Db    496 ACCACAGAAGGAACAGATTCTGGT-----------------------TTGGCTGTGAGG 531

Qy    478 LeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrp 497
Db    532 CTGGTGAATGGAGGAGACAGGTGTCGGGGTCGTGTGGAGATCCTTTACCAGGGTTCCTGG 591

Qy    498 GlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeu 517
Db    592 GGCACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAACGTGGTGTGCAGGCAGCTG 651

Qy    518 GlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyPro 537
Db    652 GGCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAGTTTGGACAGGGCTCTGGGTCC 711

Qy    538 IleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHis 557
Db    712 ATTGTTCTGGATGACGTGGCCTGTAGAGGACATGAGGCCTATCTGTGGAGCTGCTCCCAC 771

Qy    558 SerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer--- 576
Db    772 CGAGGCTGGCTCTCTCATAACTGTGGACATCAGGAGGATGCTGGAGTCATCTGTTCATAT 831

Qy    576 ------------------------------------------------------------ 576
Db    832 TCTCAAACAAGCAGTCCCACACCCGATTCTCAAACAAGCAGTCCCACACCCGGTTGGTGG 891

Qy    577 ------------------GlyAspAlaThrTrpGly------------------------ 582
Db    892 AACCCCGGGGGCACAAATAACGATGTGTCCTATGGACCCGAACAGACCACAGACGCAACA 951

Qy    583 ---------------LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGlu 597
Db    952 GATTCTGGTTTGGCTGTGAGGCTGGTGAATGGAGGAGACAGGTGTCAGGGTCGTGTGGAG 1011

Qy    598 ValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAla 617
Db   1012 ATCCTTTACCAGGGTTCCTGGGGTACCGTGTGTGACGACAGCTGGGACACCAAGGATGCC 1071

Qy    618 AlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsn 637
Db   1072 AACGTGGTGTGCAGGCAGCTGGTCTGTGGCTGGGCCTTGTCTGCCCCAGGAAGTGCCCAC 1131

Qy    638 AlaSerThrGlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
Db   1132 TTTGGACAAGGCTCTGGATCCATTGTTCTGGATGACGTGGCCTGTACAGGACATGAGGCC 1191

Qy    658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
```

```
                 |||||||||||    :::    ||||||    :::::::::|||    |||    ||||||
Db    1192 TATCTGTGGAGCTGCTCCCACCGAGGCTGGCTCTCTCATAACTGTGGCCACCATGAGGAT 1251

Qy     678 ValGlyValIleCysSerAspAlaSer--------------------AspMet------ 688
             |||||||||||||||||||                       ||||||
Db    1252 GCTGGAGTCATCTGTTCAGATGCCCAAACCCAGAGCACAACCTGGCCAGATATGTGGCCT 1311

Qy     689 -----------Glu-LeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValGl 704
                      ||| |||:::      :::||||||    ::::::      :::          |||
Db    1312 ACTACCACTCCAGAAACTACAACAGATTGGTGGACT-ACAAAATATTCTTCCTCTGTTCC 1370

Qy     704 uValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAl 724
                                :::                |||:::    |||
Db    1371 TACAACACAATTCCCCACCATA-------------GCCGATTGGTGGACAACTCCTTCTCC 1418

Qy     724 aGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer----------- 740
             |||                       ||||||    :::    :::
Db    1419 GGAATACACC--------------TGTGGAGGTTTACTGACCCTACCCTATGGGCAGTT 1463

Qy     741 -ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysTh 760
                     |||::::::                            :::|||:::    |||
Db    1464 TTCCAGCCCATACTACCCTGGAAGC-----------TATCCTAACAATGCCAGATGT-- 1509

Qy     760 rGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHi 780
                                ||||||
Db    1510 ----------------TTGTGG-------------------------------------- 1515

Qy     780 sLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAl 800
                                :::       |||:::       :::       :::|||
Db    1516 ----------------AAAATTTTCGTCTCCAGCATGAACCGTGTGACAGTGGTCTTCAC 1559

Qy     800 aAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCy 820
             |||:::             |||                                    ||
Db    1560 AGATGTGCAGCTTGAAGGAGGT---------------------------------TG 1583

Qy     820 sAspSerAspPheSerLeu------------------HisAlaAlaAsnValLeuCysAr 834
            |:::    |||:::    |||                    :::    ::::::::::::    ||
Db    1584 CAACTATGACTACATCCTGGTTTTTGATGGCCCTGAAAACAATTCTTCTCTCATTGCTCG 1643

Qy     834 gGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly-- 853
            |                                      |||    |||    |||
Db    1644 G-------------------------------GTTTGTGATGGGTTCAATGGATCTTT 1670

Qy     854 ----------------------------------------AsnGlyLeuThrTrpAlaGl 860
                                                    :::|||
Db    1671 CACCTCAACCCAGAACTTCATGTCTGTAGTCTTTATCACGGATGGCAGTGTCACGAGGAG 1730

Qy     860 uLysPheGlnCysGlu--------------------GlySerGluThrHisLeuAlaLe 873
              ||||||    :::                         |||    |||
Db    1731 AGGGTTCCAAGCTGACTACTACTCCACTCCTATCAGCACCAGCACAACCTCTCCAACGAC 1790

Qy     873 uCysProIleVal------------------GlnHisProGluAspThrCys------Il 885
              |||||||||                        ||||||    ||||||          ::
Db    1791 GTTCCCGATCGTTACTGATTGGTGGACAACTCCTTCTCCGGAATACACCTGTGGAGGTTT 1850

Qy     885 eHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGl 905
            :    :::             |||          |||    |||
Db    1851 ACTGACCCTACCCTATGGGCAGTTTTCCAGCCCATAC----------------------- 1887

Qy     905 yLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCy 925
                                                        :::    ||||||
Db    1888 ------------------------------------------TACCCTGGAAGCTAT-- 1902

Qy     925 sAspThrHisTrpAspProGluAspAlaArgValLeuCysArg----------------- 939
                                |||    :::|||||||    |||    :::
Db    1903 ----------------CCTAACAATGCCAGATGTTTGTGGAAAATTTTCGTCCCCAGCAT 1946

Qy     940 -------------------------------GlnLeuSerCysGlyThr----------- 945
                                           ||||||       |||
Db    1947 GAACCGTGTGACAGTGGTCTTCACAGATGTGCAGCTTGAAGGAGGTTGCAACTATGACTA 2006

Qy     946 -AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHi 965
                |||       |||    :::|||                        ||||||
Db    2007 CATCCTGGGTTTTGATGGTCCTGAATACAATTCTTCTCTCATTGCTCGGGTT-------- 2058

Qy     965 sArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGl 985
                       |||    |||:::                       :::    |||
Db    2059 ----------TGTGATGGGTCCAATGGATCTTTCACCTCAACCCAG------------- 2094

Qy     985 yAlaProProCysIleHisGlyAsnThrValSerVal------IleCysThrGlySerLe 1003
                                 |||    :::||||||       |||       ||||||::
Db    2095 ---------------------AACTTCATGTCTGTAGTCTTTATCACGGATGGCAGTGT 2132

Qy    1003 uThrGlnProLeuPhe--------------------ProCysLeuAlaAsnValSerAs 1016
            :|||:::    |||                    |||          :::    :::
Db    2133 CACGAGGAGAGGGTTCCAAGCTGACTACTACTCCACTCCTATCAGGACCAGCACAACTCC 2192

Qy    1016 pProTyrLeuSerAlaValProGluGly------SerAlaLeuIleCysLeuGluAspLy 1034
             |||             :::    |||       |||:::|||:::
Db    2193 TCCAACGACGTTCCCGATCATTACTGGAAATGATTCTTCATTGGTG------------- 2238

Qy    1034 sArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAs 1054
                ||||||||||||:::|||    :::||||||    ||||||||||||||    :::
Db    2239 ----CTGAGGCTGGTAAATGGAACAAACCGGTGTGAGGGCCGAGTGGAGATCTTGTACAG 2294

Qy    1054 rGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCy 1074
             |||    |||               ||||||    ||||||::::::||||||:::||||||||
Db    2295 AGGCTCTTGGGTACCGTGTGCCGACGACAGCTGGGACATCAATGATGCCAATGTGGTCTG 2354

Qy    1074 sGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGl 1094
            |::::::||||||||||||    |||    :::|||          :::|||.    ||||||:::||
Db    2355 CAGACAGCTCGGTTGTGGCTCTGCTCTGTCTGCTCCAGGAAATGCTTGGTTTGGTCAGGG 2414

Qy    1094 ySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGl 1114
            ||||||    |||    |||||||||::::::|||:::|||    |||||||||||||||
Db    2415 TTCAGGGCTCATTGTCCTGGATGATGTGTCTTGCTCTGGGTATGAGTCCCACCTGTGGAA 2474

Qy    1114 nCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIl 1134
             |||          ||||||    |||:::|||||||||    |||||||||||||||
Db    2475 TTGTCGTCACCCTGGCTGGCTTGTTCATAATTGTCGTCATGTTGAGGATGCAGGAGTCAT 2534

Qy    1134 eCysSer------GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCy 1152
            ||||||       :::    |||    |||          |||          :::::    ||
Db    2535 TTGCTCACTCCCTGATCCGACTCC-CTCTCCTGGTCCAGTTTGGACAAGTCCTCCTTTTG 2593

Qy    1152 sAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIl 1172
            |       :::|||    ::::::                |||    |||:::|||
Db    2594 T------AAACTATACTTGTGGAGGTTTCCTGACTGGAC-TCTCTGGGCAATTTTCTAGC 2646

Qy    1172 eThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValVa 1192
                |||    :::    |||:::                ||||||
Db    2647 CCATACTACCCTGGGAGCTATCCTAATAATGCCAGATGTTTGTGGA------------- 2692

Qy    1192 lSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp-------- 1209
                            |||:::|||    :::       :::|||          :::
Db    2693 -------ACATTGAAGTCCCAAACAACTACCGCGTGACTGTGGTCTTCAGAGATGTGCAG 2745

Qy    1210 -------------------------------------IleGlnCysProLysThr---Hi 1216
                                                 :::       |||    |||    ||
Db    2746 CTGGAAGGGGGCTGCAACTATGACTATATAGAGATTTTGATGGCCCCCACCACAGTTCA 2805

Qy    1216 sIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro 1234
            |:::|||:::          :::       ||||||          :::       |||
Db    2806 CCTCTCATTGCCCGGGTTTGTGATGGGGCCATGGGCTCTTTCACTTCAACATCCA 2860
```

RESULT 3
US-09-341-587-6
; Sequence 6, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096

```
;  EARLIER FILING DATE: 1998-01-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 29598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-6

Alignment Scores:
Pred. No.:              3.89e-61     Length:         29598
Score:                  785.50       Matches:        501
Percent Similarity:     26.46%       Conservative:   185
Best Local Similarity:  19.32%       Mismatches:     520
Query Match:            9.65%        Indels:         1390
DB:                     4            Gaps:           103

US-09-759-130B-381 (1-1453) x US-09-341-587-6 (1-29598)

Qy     35 CysPheLeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAsp 54
          |||   :::  |||   :::|||   |||  |||||||||||||||||||||
Db   3837 TGCCTTCCTCTAGGATCTGATTCTGGT------TTGGCCCTGAGGCTGGTGAATGGAGAT 3890

Qy     55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
          |||  |||  |||  ||||||:::  :::::::|||  |||||||||||||||||
Db   3891 GGCAGGTGTCAGGGCCGAGTGGAGATCCTATACCGAGGCTCCTGGGGCACCGTGTGTGAT 3950

Qy     75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
          |||   |||:::|||    :::  ||||||||:::||||||||||   ::::::
Db   3951 GACAGCTGGGACACCAATGATGCCAACGTGGTCTGTAGGCAGCTGGGTTGTGGCTGGGCC 4010

Qy     95 PheAlaMet---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrp 111
             :::             |||||||||              |||   |||
Db   4011 ATGTCAGCTCCAGGAAATGCCTGGTTTGGCCAGGGC---------TCAGGACCCATTGCC 4061

Qy    112 LeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGlu 131
          ||||||||||||  |||  |||::||||||  ||||||   |||   |||
Db   4062 CTGGATGATGTGCGCTGCTCAGGACACGAATCCTACCTGTGGAGCTGCCCCCACAATGGC 4121

Qy    132 TrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCysTyrGlyGluAla 151
          |||  ||||||||||||  ||||||||||||  ||||||   |||   |||
Db   4122 TGGCTCTCCCATAACTGTGGCCATGGTGAAGATGCTGGTGTTATCTGCTCAGGTAGG--- 4178

Qy    152 AsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCys--------------------- 164
          :::      |||       |||          ||||||
Db   4179 CATCCAGATCTCTGGAGGGTTGGGTGTGGTGGCTCATGCCTTTAATCCCCACACTTTGGG 4238

Qy    165 ------------------------SerGlyArgValGluValLysPheGlnGluArgTrp 176
                                  |||     |||:::   :::         ||||||
Db   4239 AGGATGAGGTAGGCAGATTGCTTGAGCTCAGGCGTTCAAGACCAGCTCTAGGCAAGATGG 4298

Qy    177 GlyThr------------------------------------------------------ 178
             |||
Db   4299 CAAACCCCATCTCTATTAAAAAAAGAAAATCACTTGGGCCCATTCTGGGGACAGACTGTA 4358

Qy    178 ------------------------------------------------------------ 178

Db   4359 TTCAGAGAATATAGTAACAGGTTTCTCACTCCAGGGCCCCCGCCCAAGTCTTTCCTGGTT 4418

Qy    179 ------------------------------------------IleCysAspAspGlyTrp 184
                                                    :::         |||
Db   4419 ATCTGGTGGCTCTGGAGTTTAGTTGGGGCAGTTGGCCTGTGGGTACAATGCCACGGTCAG 4478

Qy    185 AsnLeuAsnThrAlaAlaValValCysArgGlnLeuGly--------------------- 197
          :::  :::   :::        ||||||  :::|||
Db   4479 CACAAATCCCCAAGCTGCGAGCAGTGCCGCAACATTGGTTTGGAGGTGACTGAGGCACCA 4538

Qy    197 ------------------------------------------------------------ 197

Db   4539 GTGTTCAGACACAGGGTCCAGCATGCTGACATCAGAAATGATGAGCTTTTATTCTGCTC 4598

Qy    197 ------------------------------------------------------------ 197

Db   4599 CAGGAACATCCGCATAACTCATACCCCCATTTTCCCCGAGGCTCCTTCTCAGGGACAGCA 4658

Qy    198 ---------------------------------------------CysProSerSerPhe 202
                                                       |||   :::
Db   4659 CGTTTTGGGGATCTGGTTGGTGAAGCCTTCCCTCCCCTTGAACACTGTATTACCTTGAGT 4718

Qy    203 IleSerSer--------------------GlyValValAsnSerProAlaValLeu--- 214
          :::  |||                     :::|||||||||     ::::::
Db   4719 GTTGTCAGCCGCTGACCCAGAGGGAAGGATGCAATTGTAAATAGTTCACTTATGATTGCT 4778

Qy    215 ------------ArgProIle---------TrpLeuAspAspIleLeuCysGlnGlyAsn 227
                      |||   :::          |||   |||           |||:::
Db   4779 GTGAAGAGAGGCAGGGAAGTGGGGTAAGGGTGGGAGGAT--------------GGCAGT 4823

Qy    228 GluLeuAlaLeu-TrpAsn---------CysArgHisArgGlyTrpGlyAsnHisAspCy 244
          |||     ::: |||              :::||||||     |||:::
Db   4824 GAGTCGCAAATGGTGGGCTCACGAGTGAGCTCAGCATAGGGAGCAGGGGGACCTCCTAGA 4883

Qy    244 sSerHisAsnGluAsp-------------------------------------------- 249
           |||   |||
Db   4884 GTATCACCAGGAATCCACGTCAAACCTCAGACTCATCGCAGCTCAAGCGTGGAGAGTTAG 4943

Qy    250 ------------------------ValThrLeuThrCysTyrAspSerSerAspLeuGl 261
                                  |||   :::  |||            |||  ||
Db   4944 GGTCTTCATCTACTCATTGCTTAGGGTGGCTGTTTGCTGCTCCCTGGGTGTTGATACCGA 5003

Qy    261 uLeuArgLeuValGly-------------------------------------------- 266
          |:::   :::  |||
Db   5004 GGTTGTTGTGGCCGGTCCCAGGCACCGAAGTGACCTTCATATCCCTGGAGAGTGAGCCAC 5063

Qy    267 -GlyThrAsnArgCysMet----------------------------------------- 272
           ||||||   :::   |||
Db   5064 AGGCACAGAGAAGGCAATGCCAGGGGCTGTGGGGTCCCTGAGGGACCGAGGCTTCACGG 5123

Qy    273 ------------------------GlyArgValGlu-LeuLysIleGlnGly------- 281
                                  |||||||||      ||||||
Db   5124 TGGGCACTGGCAGGGCCCACTTGGTGGGCGTGTGATGGGCATCAGCTCAGGGTGTAGATA 5183

Qy    281 ------------------------------------------------------------ 281

Db   5184 CCCCAAGTCACTTCAGCCTTAACTCTACTTGGAGTCACTGAGTGTTTGGTGTCTAATGTT 5243

Qy    281 ------------------------------------------------------------ 281

Db   5244 GCTATTTTTTTCTCACAGCTGCCCAGCCTCAGTCAACACTCAGGCCAGGTGAGTCCCCAG 5303

Qy    281 ------------------------------------------------------------ 281

Db   5304 AATCCTTCCTCGGGATACCCCTTCTCTTTCTGCTCAGTTACCCCTTCCCTACTCCACAGA 5363

Qy    282 ----------------------ArgTrpGlyThrValCysHis---------------- 288
                                |||||    :::    |||
Db   5364 GCCCTCCTGCTTCTCTGCAGATACTCTGGGGCATATTATTCACCCCCAACTCTGTAACT 5423

Qy    289 -------------------------HisLysTrpAsnAsnAlaAlaAla---------- 296
                                   |||  |||:::   |||:::|||
Db   5424 GAGACCCTAGCATGGGGCTTCTTAACCACACATGGGATTTAGCTTCTGCCTTCTCTTCAG 5483

Qy    297 -------------------AspValValCys---------------------------- 300
                             |||   |||
Db   5484 TCCATCTCAGCTTTCATCACACAGTTCCATGCTGTCCCAATGAACAACTCTTAGAAGTTC 5543

Qy    301 --LysGlnLeuGlyCys------GlyThrAlaLeuHis-PheAlaGlyLeu--------- 314
            |||:::     |||             |||    ::::::||||||
Db   5544 AGAAAGAAGCCTGCTGTTTAATGAGCTTTGGCCTCTTTATATTCAGGGCTGACACAACCT 5603

Qy    314 ------------------------------------------------------------ 314

Db   5604 TTCTGGGAATTGAGGGTGAGACTCTCTAATGTTCTGTTGAAGGCAAGGTCAGTACAGGCT 5663

Qy    315 ------ProHisLeu--------------------------------------------- 317
                |||||||||
Db   5664 TGATACCCCCATCTCTCACTACTTGAAAAACCCCCAGATTCTGCAGGGCCACGTCTGCAT 5723
```

```
Qy    317 ------------------------------------------------------------ 317

Db   5724 TCAGAAGAGGCAGAGGCCATGCTGAGGGGAGAGAGAGGCTAAGAAATGTTTCTGGTGCCT 5783

Qy    317 ------------------------------------------------------------ 317

Db   5784 TCACTTACCAGGAAACCTGATTTCCCCCAGGGCGGACCCCTAGTATCCCGAACATTTTAG 5843

Qy    318 -----------GlnSerGlySerAspVal------------------------------- 323
                       ||||||   |||
Db   5844 CTGCAAGTGTCAAGTCTTGGCAGTGGTGTCAGATGAGCCAGTCAGTCCACGCGTCAGTGG 5903

Qy    323 ------------------------------------------------------------ 323

Db   5904 ATGGTGGGGGATGGGGGTGGTGAAGGTTTCCTAACTGCGAATAGGTCATCCCTCTCTCAG 5963

Qy    323 ------------------------------------------------------------ 323

Db   5964 AGAGAGGTGGAAGGGCCTGCATGGTGTCCTTTGTCCCTGAATGAGTTCCCTGGGCAGGAG 6023

Qy    324 ValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHis 343
             |||   |||                ||||||   :::::::|||   |||:::|||   |||
Db   6024 ACCTGGGCAGACACATGGGGAGCAAGTGGC---CAGACCTTCGAGTGGAATTGTTTTCAC 6080

Qy    344 SerGlyThrValAsnPheAspCysLeuHisGlnAsnAspVal------------------ 357
          |||                        ||||||   |||
Db   6081 AGTGCT--------------TGCCTGGTCCAGAGAACCGCTTGTTTTTACCTTTTTCC 6125

Qy    358 ---------------------------------------------SerValIleCysSer 362
                                                       |||      |||
Db   6126 CTTCAAGTCCAATTGTATCCTTTCTCTTTGTTGCTGTTTACAGAAAGTTG-GCCTGTCAG 6184

Qy    363 Asp------------------------------------------------GlyAlaAsp 366
          |||                                                ||||||
Db   6185 GATATCACCACCTGTACCCACAGAAGGTAAAGAATCCTCTCAACACTCCCTGGGGCTCAC 6244

Qy    366 ------------------------------------------------------------ 366

Db   6245 TTTCTACCTCTGGATACACTTTGGATTTCATAGTTCACTTATGATTGCCATAAAGAGAGG 6304

Qy    367 ---------------------------------------------------LeuGluLeu 369
                                                             |||:::
Db   6305 TGGGGAAGTGGGCTAAGCGTGGGAGGGTGGCAGCAGGTGATAAACAGTTGGCTCAAGAGT 6364

Qy    370 ArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHis------ 387
          :::|||         |||:::   :::   :::   :::         ::::::   |||
Db   6365 CAGCTCGGCACTGGGAACAAGAGCAGGACCTCCCAGAAGATCATTAAGAATCATTAGGCT 6424

Qy    387 ------------------------------------------------------------ 387

Db   6425 CATGATAGGATGAGGCTCAAGGTGGGCCCCTTGCTTTTTCATGTTTCTGTGGGTTGGGTA 6484

Qy    388 -----------GluGln-------------------------TrpTrpThrIle------ 393
                     |||                          :::|||   :::
Db   6485 GGGAGGAAGCTGGAGTCTCTGAAAACCCAGAATTAGATGTGATATTGGAGGGTGGAGGGT 6544

Qy    394 ---------------------------CysAspGlnAsnTrpLysAsn------------ 400
                                     |||   :::   |||
Db   6545 GCTGGTGACCTGTCTCCTGTGGGATCCTGTTCCAAG---TGGTCAGAAAAGATCCTTATC 6601

Qy    400 ------------------------------------------------------------ 400

Db   6602 TATGGGCTCAGAACAAGCCCTGGGGGTCTCCCTAATCCTATGGGGACCTCATCCCTGCCA 6661

Qy    400 ------------------------------------------------------------ 400

Db   6662 TCTCTGGTACTCTGCCAGTCAGATCCCAGATAGGACCATGCTCCTGAATAAGGGAGGGGT 6721

Qy    401 ---------------------GluGlnAlaLeuValValCysLysGlnLeuGly------ 411
                               ||||||         :::|||   |||||||||
Db   6722 CTGGGCCTGCCATCTGGAGCTGAGCAGCTCCATCCTCTGTGTACCCAACTGGGGAGTGGG 6781

Qy    412 ------CysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArg 429
                |||   :::         |||         |||   |||      |||
Db   6782 GCATCCATTCCCATCACATCCACTGGGGTCACAGGTGCTTCCCCAAAAGTTGAGCATCCA 6841

Qy    430 AspIleTrp---------IleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrp 446
                |||         :::   |||   :::   |||         :::   :::||||||
Db   6842 TAGACCTGGGCAGAGTAGGTTATCAGTGCTACTTTCACGGTGATGAAGCCTAGTCTGTGG 6901

Qy    447 AspCysThrTyrAspGlyLysAlaLysArgThrCys------------------------ 458
                                ||||||   |||
Db   6902 ---------------TCATATGCAAAGGTGACTGCCTGCCTAGGTGACTTAGTTCATTAG 6946

Qy    459 ------------------------------------------PheArgArgSerAspAla 464
                                                               |||
Db   6947 GAAGTACCCTGAGTGTGGAACTTACCTTAGATTCTTGACCTCATGATAGGGATGGATGAA 7006

Qy    465 GlyValIleCysSer---------AspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
          |||   :::||||||                  ::::::   |||   ||||||||||||
Db   7007 GGGTTCTTGTGTTCCCCTGTAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGA 7066

Qy    482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
                |||   ||||||:::||||||         |||:::||||||   |||||||||||||||
Db   7067 GGCGACAGGTGTCGAGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGCACCGTGTGT 7126

Qy    502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
             |||   |||   |||   :::|||   |||||||||:::||||||||||||||||
Db   7127 GATGACTACTGGGACACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGG 7186

Qy    522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
             |||         |||         |||   :::   ||||||||||||   ||||||
Db   7187 GCCATGTCAGCCCCAGGAAATGCCCAGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGAT 7246

Qy    542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
          ||||||   |||   |||:::||||||   :::|||   |||   |||:::||||||
Db   7247 GATGTGCGCTGCTCAGGACACGAGTCCTACCTGTGGAGCTGCCCCCACAATGGCTGGCTC 7306

Qy    562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGly------AspAla 579
             |||||||||   |||   ||||||         |||   |||||||||
Db   7307 ACCCACAACTGTGGCCATAGTGAAGACGCTGGTGTCATCTGCTCAGGTGGGCCTTCAAGA 7366

Qy    580 ThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArg------------ 595
          ||||||         |||   ||||||         ||||||   |||
Db   7367 ACTTGGGCTCACTCTCTTGG-GGTGGAGTT------TGCTCCAAAAGAAACTCCTAATTA 7419

Qy    596 ---------------------------LeuGluValTyrPheGlnGlyArgTrpGlyThr 606
                                               |||
Db   7420 CATTCTGATCTCCTCACTCAAAGCTTCTATGTTTCTATGTTTCTGAAGACTTGTCAGCT 7479

Qy    607 ValCysAspAspGly---------------------------TrpAsn------------ 613
          :::|||   :::                              |||:::
Db   7480 CTCTGCTAAGAATCCATATGTACTCACTGCCTAGTGTTCCTGTGGTCACTTAGGACAGGG 7539

Qy    614 ------------------------------------------SerLysAlaAlaValVal 620
                                                        ||||||
Db   7540 GATCAAACTAAAACAACCCAGAGTTTTTCCCCTTCCTGAGGCAAGGCAAGGAAGAGGCAG 7599

Qy    621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
                ::::::      |||         |||         |||   :::
Db   7600 AAGAGAAAAGTGCCGGCTCCCCAGGGCTCCATTTCTCCCCTGCTGAGTAGCACGGTGTGA 7659

Qy    641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAsp--------------- 655
          ||||||                     |||      :::   :::|||
Db   7660 GGGTATAATGGATGCAGCACAGACAGCAAGGCAGGGGAGGGATCCCCTCACTGCGAGGAA 7719

Qy    656 ------------------------------------------------GluSerAspLeu 659
                                                             ::::::
Db   7720 CTCTGAACTAAAGATGCTTGGCTAAAAGTGGGTTCTCAGCTGAGACCCAGTGAGGAGGTC 7779

Qy    660 TrpSerCysArgAsnSerGly------------Trp------------------------ 667
          |||   |||   |||                  |||
Db   7780 TGGAAATAGAGGCTCAAGGGTTAGGAGTGCAAATGGGTGTCTGTTTGTATCAGGCCTGGG 7839

Qy    668 ------------GlyAsnAsnAspCysSerHisSerGluAspValGly------ValIle 681
```

```
                           |||      |||    |||    |||::::::                     :::
Db   7840 TTGCGTGGGGTTGGAGTTCTTGACCTCAGCTCTTCTCAGAACGCTGCTGAGCATTGCCTG 7899

Qy    682 CysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGly 701
          ||||||   :::   :::                                        |||
Db   7900 TGTTCTAGGTCTGGTGAG-------------------------------------GGG 7920

Qy    702 LysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMet 721
          :::                                 |||::::::   |||     ||||||:::
Db   7921 AGGGCAGTCCCCATGAGGCCAGCCAGACATGGCCTTGTCATTGCCTGTGATTGGGGCTTG 7980

Qy    722 AsnIleAlaGlu---------------ValValCysArgGlnLeuGluCysGlySerAla 736
             ||||||:::                 |||             |||
Db   7981 AAGATCGCACAAGGGATTTTGGCTGGAGTGGCTTCCTCAGCCTTGCTGACTCAGGAACAC 8040

Qy    737 IleArgValSerArgGlu-----------------ProHisPheThrGluArgThrLeu 750
          :::|||   :::||||||                   |||   :::    :::|||
Db   8041 CTAAGATGTGCAAGGGAGTGGGTTGGTTTAGGTCAACCGGGTTACCCTGGGCAGACACAA 8100

Qy    751 ------HisIleLeuMetSerAsnSerGly-----------------CysThr-GlyGl 762
                |||:::        |||||||||                   ||| |||
Db   8101 TTTGATCACCTCAGAGCTGGCAATAGTGGACAGGATCTGCCTCGACCCCTTACACGGT-- 8158

Qy    762 yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs 782
              |||||||||||   |||        |||:::       :::   |||      ||
Db   8159 ----GCATCTCTGTGGGGATGT---GCATGGCAATGT---CCCTCCCTGTGTGATAGGAA 8208

Qy    782 nMetGlu-----------------------------------------------AlaSe 786
           |    :::
Db   8209 CTAGGATGGACTGAGTGTCAGACTCGCTCATTTCTTTCCCTCCTCGTTCCAGTTTTGCCG 8268

Qy    786 rLeuIleCysSer-------------------------------------------- 790
            |||:::|||:::
Db   8269 ACTTCTGTGTAATGTTCCTGATCTGACCTTCTCTTCTCTTTCTCACAGCTCCCCAGTCCC 8328

Qy    791 -AlaHisArgGlnProArgLeuValGlyAlaAspMetPro------------------- 803
            |||       |||||||||   |||    :::   :::|||
Db   8329 GGCCGACACCCAGCCCAGGTAGGTCCCCAGTGTCCTTCCTCAAAATGTCCCTTCTCTTTC 8388

Qy    803 ------------------------------------------------------------ 803

Db   8389 TGCCCAATCACCCCTTCCACACTCCACAGAGCTCTCCTGTTTCTCTGTGTGGATACTGTG 8448

Qy    803 ------------------------------------------------------------ 803

Db   8449 GGGCATATTATTTCCACCCCCAACACCGGCTGTGTAACTGAGATCCCAGCACAGCGCTTT 8508

Qy    803 ------------------------------------------------------------ 803

Db   8509 TTAAACACACACAGGATTGAGGAGGCCTCTGTCTTCTTTTCAACCCCTCTCAGCTTTCAT 8568

Qy    803 ------------------------------------------------------------ 803

Db   8569 GAAACAGTTTCATACTGTCCCAGTGGACAACCCTTACAGGTTCAGGAAGTGGCCCCATGT 8628

Qy    803 ------------------------------------------------------------ 803

Db   8629 TTAATGAGCTTTGGTCCTTTTATATTCCGGACTCACATATAGTTTCTGAAAATTGAGGGT 8688

Qy    804 ------------------------CysSerGlyArgValGluValLysHis-------- 812
                                     ||||||:::                |||
Db   8689 GTCACCCTCTGACCCGTTCAAGGCATCGTCAGGGCAGGCTCGATACCCCCATCCATCACT 8748

Qy    813 -AlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLe 832
              :::|||:::|||         ||||||:::   |||:::                ||
Db   8749 GCTGGAAACATTCCGAGATTAT---GATTCCGAGATTTCTATCCAGAAGAGGCAGAGTTT 8805

Qy    832 uCys------ArgGluLeu---AsnCysGlyAspAlaIleSerLeu------------- 844
           ||||     ||||||     |||    |||          :::||||||
Db   8806 GTGCTCGGGCAGGGAGAGGGATAATAAAGGTTT-GTGATGTCTCTGCTTAACCAGAAACC 8864

Qy    844 ------------------------------------------------------------ 844

Db   8865 TGATTCCTGATTGTCCCCTGGGCAGCCCCTGGTTCCCCTAACATTTTACCTGGCAGTGTC 8924

Qy    845 ------------------------------------------------SerValGlyAs 848
                                                            ||||||
Db   8925 CGAGCTTCAGCAATGGCGTCTGATGTCCTAGTCAGTCCATGCATCAGCAGATGTGGGATG 8984

Qy    848 pHisPheGlyLysGlyAsnGlyLeu----------------------------------- 856
           |||           |||    |||:::
Db   8985 GCAT---------GGTGGGGGCATCCTCTAACAGATAGAAAGATACCCCAGGATTACAGA 9035

Qy    856 ------------------------------------------------------------ 856

Db   9036 AATGGAGGGCTCAGTCTGGTCTCCAGCAGGACCTTTGTCCCTGGATGAGTTCAGCACA 9095

Qy    857 ----ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCys-P 875
               ||||||   :::   ::::::    |||                    |||
Db   9096 GGAGACCTGGGAAGACACATGGGAAACAAATGGCAGGAACAAGAAGTGGAATTGTTGTCA 9155

Qy    875 roIleValGlnHisProGluAsp------------------------------------ 882
            ::::::     |||   |||
Db   9156 CGTTGATTCCTCCTCCCAGAGATCCTTTTGTTCTGTGCCTTTTCCCTTCAAGTCTAATTC 9215

Qy    883 --------------------------------------------ThrCysIleHisSerA 888
                                                      |||||||||:::||||
Db   9216 TGTCTTTTCCTTTTGTTGCAATTTACAGATACTTGGCCGACCTCACATGCATCAACAGCA 9275

Qy    888 rgGluVal---------GlyValValCysSerArgTyr--------------------- 897
           ||   :::               :::   :::   |||
Db   9276 GGTAAATAACCCTCTCACCCCTCCCTAGGACTCACTATCTCTGGACATATTTTGTGTTTG 9335

Qy    898 --ThrAsp---ValArgLeuValAsnGlyLysSer------------------------ 907
            ||||||    :::||||||      |||   |||
Db   9336 AAACTGATAGGATGAGGCTCAATGTGGGCTTCTCTGTTTTCATGTCCCTGTGGGTTGCGT 9395

Qy    908 --GlnCysAspGlyGlnValGluIleAsnValLeuGly--------------------- 919
            :::      |||      |||      |||||||||
Db   9396 GGGAGGAAGGTGGAATCTCTGAGGAGCCAGTCCTGGGTCTGATGTTTGAGGACGGAGGGT 9455

Qy    920 -----------------------------------------------------HisTrpG 922
                                                               |||    |
Db   9456 GTTGGTGACCTGTCTCCCATGGAATCCTGTTCCAAGTGGTCAGGAAAGATCCTCATCCAG 9515

Qy    922 lySerLeuCysAspThrHisTrp---------------------------------AspP 931
           ||:::      |||   ||||||                                  |
Db   9516 GTGCTCAG---GACGAGCACTGGAGGGCTCCTTAATGCTGCTGGGACCTCATTCCTGGCC 9572

Qy    931 roGlu------AspAlaArgValLeu---------------CysArgGlnLeuSer---C 943
           ||:::        :::|||   |||                    ||||||   |||    |
Db   9573 CTCAGGCCATGGGATCCAGACCTCTGAAGAGCGGGTGAAAGTGCCGGTCCCTGCACCTGT 9632

Qy    943 ysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly-GluArgSerValArgVal 962
           |||||      |||:::   ||||||                |||  ||||||    ::::::|||
Db   9633 GTGGCCAAGGCCTTGCCATCACTGGCAATTTGCCAGAAGGCAGAGAGGCCCATGCAGGTG 9692

Qy    963 ------------------TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
                            |||                    |||:::     ::::::
Db   9693 CCAATAAGCTCCTGAATATGG------------------AGGGGGTCTAGGCCTGTCATT 9734

Qy    977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHis-------------- 991
             :::|||                 ||||||   |||:::|||
Db   9735 TGTAGCTGTGTA---------------GCTCCATCCTGTGTTCACCCAGAGTGGGGAGTG 9779

Qy    992 ------------------GlyAsnThrValSerValIleCysThrGlySerLeuThr--- 1004
                                     |||    :::   |||    ||||||:::
Db   9780 GGGTGTCCATTCCTGTCCTCTCCTCTGGGGTCATACATGCTTACCCATAGCTTGAGCTTT 9839

Qy   1005 ---------------------GlnProLeuPheProCysLeuAlaAsnValSerAsp--- 1016
                                    |||||||||          :::::::::
Db   9840 TATAGACTTGAGTAGAATAGGGCATCACTTTTTCCACTATGACCAAGCTTAACCTCTGGG 9899

Qy   1016 ------------------------------------------------------------ 1016

Db   9900 TGCAGCCATCTGCCATTGTGACTGCGTGCCCTGGTGACTTTGGCATTTGTATTCAAACTT 9959
```

```
Qy   1017 ProTyrLeuSerAlaValProGlu---------------------------------- 1024
             |||   |||    ||||||
Db   9960 GACTACTTTTCCCATTGCCCTGAGTGTTGTCCATGCCTTTTCCTTCACCTCAGAATGGAG 10019

Qy   1025 ---GlySerAlaLeuIleCysLeu--------------GluAspLysArgLeuArgLeu 1038
                              :::|||                          ||||||||
Db  10020 ATGGATGAAGGATTCTTGTGTTCCCCTGTAGGACCTGAATCCAGTTTGGCCCTGAGGCTG 10079

Qy   1039 ValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGly 1058
          |||:::|||      ||||||  ||||||||||||:::    :::   |||   ||||||
Db  10080 GTGAATGGAGGTGACAGGTGTCAGGGCCGAGTGGAGGTCCTATACCGAGGCTCCTGGGGC 10139

Qy   1059 ThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGly 1078
          |||:::|||||||||  |||:::  ||||||  |||||||||:::||||||||:::::||||||
Db  10140 ACCGTGTGTGATGATAGCTGGGACACCAGTGACGCCAATGTGGTCTGCCGGCAGCTGGGC 10199

Qy   1079 CysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIle 1098
          ||||||  |||   :::|||       :::|||  ||||||:::|||||||||||||||
Db  10200 TGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCAGGACCCATT 10259

Qy   1099 TrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArg 1118
             |||||||||:::   |||:::|||   ||||||:::||||||   ||||||
Db  10260 GTCCTGGATGACGTGCGCTGCTCAGGCTATGAGTCCTACCTGTGGAGCTGCCCCCACAAT 10319

Qy   1119 GlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer------ 1136
          ||||||     |||:::|||:::|||   ||||||||||||||||||||||
Db  10320 GGCTGGCTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCATCTGCTCAGGTGGG 10379

Qy   1136 ------------------------------------------------------------ 1136

Db  10380 CCTCCAAGACCTTGGGCTCCCTCTCTTGGGGTAGATTTTGCTCAGGAAGCGAGGTCTCAT 10439

Qy   1137 -----------------------GluPheThrAlaLeuArgLeuTyr------------ 1144
                                   |||:::|||      :::|||
Db  10440 TATGTTCTGATCTCCTCACTCAGAGCTTTTCAGCCTTTCCTATATATCTGATATCTCCT 10499

Qy   1145 -----------SerGluThrGluThrGluSer-Cys----------------------- 1152
                              :::          ||| |||
Db  10500 TAGCTCTCTCCTAGGAAACTGCATGAGTCTTCATTGCCAGGTTTTGAGGAGGTCAGGTAG 10559

Qy   1153 ----AlaGlyArgLeuGluVal------------------PheTyrAsnGlyThrTrpGl 1165
                  |||:::   |||:::                    |||   |||:::
Db  10560 GACAACGGGCCAAAGTGAAATAAGGGTCACGCCTTTGTTCACCTACCGAGGCAGC---GC 10616

Qy   1165 ySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGl 1185
           |||   |||||||||                 |||
Db  10617 AAGCAGAGGGAGAAGAGGAAAGTGCCAGGTCTTTGCC------------------TTTTA 10658

Qy   1185 yCysGlyGluAsnGlyValValSer------------LeuAlaProLeuSerLysThrGl 1201
           ||||||   |||:::                    :::       :::       ||
Db  10659 GTGTGGCTGGAAAGGAATAGCTGGGGCTGGTTTGTTCATGAGGAAAGAGGCTGATTTGGG 10718

Qy   1201 ySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGl 1221
          |  ||||||
Db  10719 TCTTGGCTTTGCA---------------------------------------------GG 10733

Qy   1221 nCysLeuSerAlaProTrpGluArgArgIleSerSer---------------------- 1233
           |||:::     ||||||  ||||||:::   |||
Db  10734 CTGCATA---GGACCATGGTGCAGGCGTCTGCTCAGCTTCTGATGAGGGCCTCGGGCTGC 10790

Qy   1234 -------ProAlaGluGluThrTrp---------IleThrCysGluAspArgIleArgVa 1248
                 |||:::   :::   |||          :::          ::::::|||
Db  10791 TTGTACTCCTAGCAGAAAGGGATGGGGAGGTGGCCTGGGCAGAGGTCACAAGGTGAGGGA 10850

Qy   1248 lArgGlyGlyAspThrGluCysSerGly-------------------------------- 1257
            |||   :::::::::       |||
Db  10851 GGAAGGAAGAGAGAGCAAGAGGGAGGGCTCAGACTCTTCAACCACCAGCTCTTGCTAGGA 10910

Qy   1257 ------------------------------------------------------------ 1257

Db  10911 ACTAAGAGAAGAACTCACCCCTGACCAGGGAGGGCACTAAGTTATTCATGCAGTGTTGGT 10970

Qy   1258 ----ArgValGluIleTrp------------HisAlaGlySerTrpGlyThrValCysAs 1272
                  :::   |||             |||      :::||||||      |||::
Db  10971 CTCCATGACCCAGACCTGGTGCATCAGGCCTCACCTCTAAACTTGGGGATTCAGTTGCAA 11030

Qy   1272 pAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys-------- 1289
          :    :::                                   |||||||||
Db  11031 TATAACA-----------------------------------CTTGGATGTGACAAGCC 11054

Qy   1290 ----------------------------GlySerAlaLeuAlaAlaLeuArgAspAlaSe 1300
                                      ||||||       |||:::|||
Db  11055 TCTAAAGTATAGCACCCTGGCTCTCCAGGGTTCCATCTTTCCCCTACTGAAAGAT----- 11109

Qy   1300 rPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSe 1320
            |||    :::|||             ::::::|||:::
Db  11110 -TTTGCAAAGGGT------------TACATGGATGAA---------------------- 11133

Qy   1320 rPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAs 1340
                                                        |||   |||||
Db  11134 ---------------------------------------------GCACATATAGAGGA 11147

Qy   1340 pAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLe 1360
          |   |||                           ||||||       :::      |||
Db  11148 TGGGGGA------------------GGGACTGTCTCCCTGGGAGGAGCCCAGAACCACAG 11189

Qy   1360 uAlaLeuIleLeuSerSerIlePhe 1368
              |||:::   |||::::::|||
Db  11190 GCTCCTGCTAGGAAGCGCGGTCTTC 11214
```

RESULT 4
US-09-341-587-2
; Sequence 2, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-2

```
Alignment Scores:
Pred. No.:                1.51e-56      Length:          2001
Score:                    715.50        Matches:         164
Percent Similarity:       48.18%        Conservative:    61
Best Local Similarity:    35.12%        Mismatches:      119
Query Match:              8.79%         Indels:          123
DB:                       4             Gaps:            15

US-09-759-130B-381 (1-1453) x US-09-341-587-2 (1-2001)

Qy    956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975
          |||      |||      :::|||                  |||      |||      |||
Db      1 GGACACGAGTCTTAWABCCTGTGGAGC------------TGCCCCCACAATGGCTGGCTC 48

Qy    976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrVal 995
                |||||                                |||
Db     49 TCCCACAACTGTGGC-----------------------------CATCATGAAGATGCT 78

Qy    996 SerValIleCysThrGlySerLeuThrGln------ProLeuPheProCysLeuAlaAsn 1013
             |||||||||:::   :::   :::|||      |||
Db     79 GGTGTCATCTGCTCAGCTGCTCAGTCCCAGTCAACGCCCAGGCCAGATACTTGGCTGACC 138

Qy   1014 ValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAsp 1033
             :::   |||   |||:::             |||   |||
Db    139 ACCAACTTACCGGCATTGACAGTAGGATCTGAATCCAGTTTG------------------ 180
```

```
Qy   1034 LysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHis 1053
             ||||||||||||:::|||        ||||||  ||||||||||||:::    :::
Db    181 ---GCTCTGAGGCTGGTGAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTCCTGTAT 237

Qy   1054 AspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValVal 1073
             |||   |||||||||:::|||||||||     ||||||    :::|||||||:::||||||
Db    238 CGAGGCTCCTGGGGAACCGTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTC 297

Qy   1074 CysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGlu 1093
          |||::::::||||||||||||   |||   :::|||      :::|||   ||||||:::
Db    298 TGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCGGCCCCAGGAAATGCCCGGTTTGGCCAG 357

Qy   1094 GlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrp 1113
          |||||||||||||||      |||||||||:::     |||:::|||   ||||||:::||||||
Db    358 GGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGGAATGAGTCCTACCTGTGG 417

Qy   1114 GlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyVal 1133
             ||||||   :::||||||      |||:::|||   |||   |||||||||||||||
Db    418 AGCTGCCCCCACAAAGGCTGGCTCACCCACAACTGTGGCCATCACGAAGACGCTGGTGTC 477

Qy   1134 IleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThr------------------ 1147
          |||||||||         |||   ::::::   |||   |||
Db    478 ATCTGCTCA---------GCCACCCAAATAAATTCTACTACGACAGATTGGTGGCATCCA 528

Qy   1148 --------------------GluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
                              :::    :::|||   |||   |||      ||||||
Db    529 ACAACTACAACCACTGCAAGACCCTCTTCAAATTGTGGTGGCTTCTTA------TTCTAT 582

Qy   1161 ---AsnGlyThrTrpGlySer----------------------------------- 1166
             :::|||||||:::   |||
Db    583 GCCAGTGGGACATTCTCCAGCCCATCCTACCCTGCATACTACCCCAACAATGCTAAGTGT 642

Qy   1167 ---------------------------------------ValGlyArgArgAsnIle--- 1172
                                                 :::|||      |||:::
Db    643 GTTTGGGAAATAGAAGTGAATTCTGGTTATCGCATAAACCTGGGCTTCAGTAATCTGAAA 702

Qy   1172 ------------------------------------------------------------ 1172

Db    703 TTGGAGGCACACCATAACTGCAGTTTTGATTATGTTGAAATCTTTGATGGATCATTGAAT 762

Qy   1173 ThrThrAlaIleAlaGlyIleValCys---------ArgGlnLeuGlyCysGlyGluAsn 1189
          ::::::   :::   |||   :::|||          ||||||::x
Db    763 AGCAGTCTCCTGCTGGGGAAAATCTGTAATGATACCAGGCAAATATTTACATCTTCTTAC 822

Qy   1190 GlyValValSerLeu---------AlaProLeuSerLysThrGlySerGlyPheMet--- 1205
            :::::::::           :::   :::|||            :::|||||||:::
Db    823 AACCGAATGACCATTCACTTTCGAAGTGACATCAGTTTCCAAAACACTGGCTTTTTGGCT 882

Qy  .1206 TrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAla 1225
          |||   :::
Db    883 TGGTATAAC--------------------------------------------------- 891

Qy   1226 ProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArg 1245
                            |||   |||::::::   |||
Db    892 -----------------TCCTTCCCAAGCGATGCCACCTTG-----------AGGTTG 921

Qy   1246 IleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGly 1265
          :::   :::                     |||:::|||||||||||||||:::|||   |||
Db    922 GTCAATTTAAATTCATCCTATGGTCTATGTGCCGGGCGTGTAGAAATTTACCATGGTGGC 981

Qy   1266 SerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGln 1285
          :::||||||||||||||||||||||||||||||      :::   ||||||||||||||||||:::
Db    982 ACCTGGGGGACAGTTTGTGATGACTCCTGGACCATTCAGGAAGCTGAGGTGGTCTGCAGA 1041

Qy   1286 GlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThr 1305
          ||||||||||||||     |||::::::||||||    :::|||    ||||||    |||:::
Db   1042 CAGCTAGGGTGTGGACGTGCAGTTTCAGCCCTTGGAAATGCATATTTTGGCTCTGGCTCT 1101

Qy   1306 GlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCys 1325
          |||   |||   |||||||||:::   |||   |||   ||||||   ||||||   .|||
Db   1102 GGCCCCATCACCCTGGACGATGTAGAGTGCTCAGGGACGGAATCCACTCTCTGGCAGTGC 1161

Qy   1326 HisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCys 1345
                 :::   |||        :::|||   |||:::||||||||||||||||   |||
Db   1162 CGGAACCGAGGCTGGTTCTCCCACAACTGTAATCATCGTGAAGATGCTGGTGTCATCTGC 1221

Qy   1346 SerGlyGlnSerLeuLysSer 1352
             ||||||    |||   :::
Db   1222 TCAGGAAACCATCTATCGACA 1242
```

RESULT 5
US-09-034-916-1
; Sequence 1, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
;   APPLICANT: GEBE, JOHN A.
;   APPLICANT: SIADAK, ANTHONY W.
;   APPLICANT: ARUFFO, ALEJANDRO A.
;   TITLE OF INVENTION: SpALPHA: A NOVEL SCAVENGER RECEPTOR
;   TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
;   TITLE OF INVENTION: ANTIBODIES THERETO
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ROBINS & ASSOCIATES
;   STREET: 90 MIDDLEFIELD ROAD, SUITE 200
;   CITY: MENLO PARK
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/034,916
;   FILING DATE: 04-MAR-1998
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/039,956
;   FILING DATE: 06-MAR-1997
;   ATTORNEY/AGENT INFORMATION:
;   NAME: ROBINS, ROBERTA L.
;   REGISTRATION NUMBER: 33,208
;   REFERENCE/DOCKET NUMBER: 5998-0020
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 325-7812
;   TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2178 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 64..1101
US-09-034-916-1

Alignment Scores:			
Pred. No.:	2.94e-54	Length:	2178
Score:	691.50	Matches:	148
Percent Similarity:	52.51%	Conservative:	40
Best Local Similarity:	41.34%	Mismatches:	139
Query Match:	8.50%	Indels:	31
DB:	3	Gaps:	9

US-09-759-130B-381 (1-1453) x US-09-034-916-1 (1-2178)

```
Qy     21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----------Phe 36
             |||||||||                  |||:::|||      |||            |||
Db     70 CTATTCTCC-----------------TTGATCCTTGCCATTTGCACCAGACCTGGATTC 111

Qy     37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
```

```
                 |||    |||                               :::|||||||||    |||
Db     112 CTAGCGTCTCCATCTGGA--------------GTGCGGCTGGTGGGGGGCCTCCACCGC 156

Qy      57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
           |||    |||    |||||||||:::     :::||||||||||||||||||||||||||||
Db     157 TGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATGACGGC 216

Qy      77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
           |||:::                   |||:::|||:::::||||||||         :::
Db     217 TGGGACATTAAGGACGTGGCTGTGTTGTGCCCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 276

Qy      97 Met---------PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
                                     :::               :::      |||:::    :::
Db     277 ACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCAA 336

Qy     114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
              ||||||||    |||    |||        |||    :::|||:::       |||
Db     337 TCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAA------ 390

Qy     134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------------------ 147
              ::::::|||    |||    ||||||      |||    :::|||
Db     391 GTTTATGATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCT 450

Qy     148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
           :::                |||:::||||||      ||||||            |||    |||    ||||||
Db     451 TTCTCCCCAGTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGC 510

Qy     168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
           ||||||||||||    |||       :::|||    |||:::|||       ||||||:::|||
Db     511 GTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGG 570

Qy     188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
              |||    ||||||||||||||||||||||||              :::    ::::::
Db     571 GCCGCAAAGGTGGTGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGC 630

Qy     208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
              |||         |||        |||    |||||||||||||         :::    |||    |||
Db     631 TGCAACAAGCATGCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 690

Qy     227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
              |||         |||       :::|||            ||||||      :::    |||:::|||
Db     691 CGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCAT 750

Qy     247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
           :::||||||        .:::    |||    |||         |||         |||||||||||||||
Db     751 GATGAAGACACGTGGGTCGAATGTGAAGATCCCTTTGAC------TTGAGACTAGTAGGA 804

Qy     267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
           |||    |||    |||    ||||||:::|||:::         :::|||    ||||||:::|||
Db     805 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGTC 864

Qy     287 CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
           |||            |||                     |||||||||||||||||||||||||||
Db     865 TGTGATGACAACTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGGG 924

Qy     307 ThrAlaLeu------HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
              :::|||              |||                        |||           :::
Db     925 AAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCATC 984

Qy     325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
           |||||||||    |||    |||||||||    |||         |||         |||:::|||
Db     985 TGGCTGGATAATGTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAGA 1044

Qy     345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
                                ||||||    ||||||    ||||||:::||||||||||||
Db    1045 TTTTGGGGGTTTCACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCA 1098
```

RESULT 6
US-09-276-400-3
; Sequence 3, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-276-400-3

```
Alignment Scores:
Pred. No.:                 6.12e-49        Length:        2262
Score:                     634.50          Matches:       214
Percent Similarity:        40.03%          Conservative:  97
Best Local Similarity:     27.54%          Mismatches:    253
Query Match:               7.80%           Indels:        215
DB:                        3               Gaps:          32

US-09-759-130B-381 (1-1453) x US-09-276-400-3 (1-2262)

Qy      30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr--------------- 44
           |||||||||          |||    |||    ||||||          |||:::
Db      40 CTGCTGCTG------TGCCTGCTGTGCAGTTCGTGCTTGGGGTCTCCGTCCCCTTCCACG 93

Qy      45 ---------------------------AspLeuGluLeuArgLeuValAsn---GlyAsp 54
                                                 |||          ||||||
Db      94 GGCCCTGAGAAGAAGGCCGGGAGCCAGGGGCTTCGGTTCCGGCTGGCTGGCTTCCCCAGG 153

Qy      55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
              |||       |||    ||||||::::::          |||:::||||||||||:::||||||
Db     154 AAGCCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGAT 213

Qy      75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
           |||    :::            |||:::    ::::::|||::::::||||||         |||:::
Db     214 GATGACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGC------TTCACA 267

Qy      95 PheAla---------------MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
              |||                                       ::::::|||          |||:::
Db     268 GAGGCCACAGGCTGGACCCACAGTGCCAAATATGGCCCTGGAACA---------GGCCGC 318

Qy     110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
           ||||||||||||::::::||||||    |||    |||      ::::::   .||||||
Db     319 ATCTGGCTGGACAACTTGAGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCC 378

Qy     130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------ 147
           |||    ||||||:::      :::|||    |||    ||||||    ||||||    |||
Db     379 CGGGGCTGGGGGAACAGTGACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGAC 438

Qy     148 ---------------TyrGlyGluAlaAsnLeu-------------------------- 153
                          :::    ::::::|||:::
Db     439 CAGCGCCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 498

Qy     154 ---------------------------------GlyLeuArgLeuValAspGlyAsnAsn 162
                                            |||    |||       :::
Db     499 GAGGAGGTGCGAATTCGACCCGCCGTTGGGTGGGGCAGACGACCCCTGCCC--------- 549

Qy     163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
                    |||    |||||||||:::          :::    |||          :::||||||
Db     550 GTGACGGAGGGGCTGGTGGAAGTCAGGCTTCCTGACGGCTGGTCGCAAGTGTGCGACAAA 609

Qy     183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
           ||||||:::    :::    :::    ||||||||          ||||||    ||||||
Db     610 GGCTGGAGCGCCCACAACAGCCACGTGGTCTGCGGGATGCTGGGCTTCCCCAGCGAAAAG 669

Qy     202 PheIleSerSerGlyValValAsn------SerProAlaValLeuArgProIleTrp-Le 219
              ::::::    |||:::::          ||||||                    :::||| ||
Db     670 AGGGTCAACGC-GGCCTTCTACAGGCTGCTAGCCCAACGGCAGCAACACTCCTTTGGTCT 728

Qy     219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
           |       :::    |||    |||    |||        |||        |||           ::
```

```
Db     729 GCATGGGGTGGCGTGCGTGGGCACGGAGGCCCACCTCTCCCTCTGTTCCCTGGAGTTCTA 788

Qy     239 pGlyAsnHisAsp---------CysSerHisAsnGluAspValThrLeuThrCys----- 254
           :     :::|||          |||                       ::::::|||
Db     789 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGAGCTGTGTGCC 848

Qy     255 ----------TyrAspSerSer-------------------------------------- 258
                     |||   :::|||
Db     849 AGGCCCTGTCTACGCGGCATCCAGTGGCCAGAAGAAGCAACAACAGTCGAAGCCTCAGGG 908

Qy     259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
            :::      ::::||||||  ||||||   :::         ||||||||||||:::
Db     909 GGAGGCCCGTGTCCGTCTAAAGGGCGGCGCCCACCCTGGAGAGGGCCGGGTAGAAGTCCT 968

Qy     278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVa 298
                          ||||||||||||||     ||||||:::      ||||||  ||
Db     969 GAAGGCCAGCACATGGGGCACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1028

Qy     298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
           ||||||:::::||||||   |||:::|||                            :::
Db    1029 GGTGTGTCGGGAGCTGGGCTTCGGGAGTGCTCGAGAAGCTCTGAGTGGCGCTCGCATGGG 1088

Qy     318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
              |||         :::   |||        |||   |||||||||    |||       ||
Db    1089 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTCTGGACAGGAGCTCTCCCT 1148

Qy     338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
            ||||   |||   |||                     ||||||   |||      |||
Db    1149 CTGGAAGTGCCCCCACAAGAACATCACAGCTGAGGATTGTTCACATAGCCAGGATGCCGG 1208

Qy     358 rValIleCysSer---------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
             |||   |||:::                  ||||||:::      :::||||||:::   |||
Db    1209 GGTCCGGTGCAACCTACCTTACACTGGGGCAGAGACCAGGATCCGACTCAGTGGGGGCCG 1268

Qy     375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu------------GlnTrpTr 391
             :::          |||||||||||||||:::|||                 :::|||
Db    1269 CAGCCAACATGAGGGGCGAGTCGAGGTGCAAATAGGGGACCTGGGCCCCTTCGCTGGGG 1328

Qy     391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411
              ||||||        :::|||            :::|||:::|||   |||:::||||||
Db    1329 CCTCATCTGTGGGGATGACTGGGGGACCCTGGAGGCCATGGTGGCCTGTAGGCAACTGGG 1388

Qy     411 yCysProPheSerValPheGlySerArgArgAla------LysProSerAsnGluAlaAr 429
           |       ::::::        |||    :::                   |||
Db    1389 TCTGGGCTACGCCAACCACGGCCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1448

Qy     429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
            ::::::   ::::::   :::    ||||||||   |||    :::|||       |||
Db    1449 AGAGGTGGTGATGAGTGGAGTGCGCTGCACAGGGACTGAGCTGTCCCTGGATCAGTGTGC 1508

Qy     449 rTyrAspGlyLysAlaLysArgThrCysPheArg---------ArgSerAspAlaGlyVa 466
            :::   |||            ||||||   |||              |||     ||||||||
Db    1509 CCATCATGGC---ACCCACATCACCTGCAAGAGGACAGGGACCCGCTTCACTGCTGGAGT 1565

Qy     466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486
           ||||||||||:::      |||       ||||||   |||        ||||||
Db    1566 CATCTGTTCTGAGACTGCATCA---GATCTGTTGCTG---------CACTCA-------- 1605

Qy     486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
                                                  |||      ||||||
Db    1606 --------------------------------GCACTGGTGCA-GGAGACCGCCTACA 1630

Qy     506 rThrArgAsnAla---------AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
               |||  .  |||                |||:::|||           ||||||:::      ::
Db    1631 TCGAAGACCGGCCCCTGCATATGTTGTACTGTGC-----------TGCGGAAGAGAACT 1678

Qy     523 t---------------HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
           :                 :::        :::        :::      ||||||    |||||
Db    1679 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGGTCACCGGCGTCTGCTCCGAT 1738

Qy     538 eTrpLeuAspAspVal-SerCys-----------------------IleGlyAsnGluS 550
           |   |||   |||       :::                                  |||       |
Db    1739 TCTCCTCCCAGATCCACAACCTGGGACGAGCTGACTTCAGGCCCAAGGCTGGGCGCCACT 1798

Qy     550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
            ||   :::|||    :::|||                      |||          |||
Db    1799 CCTGGGTGTGGCACGAGTGC--------------------CATGGGCATTACCACAGCA 1837

Qy     569 luAspVal---------------IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
             |||:::                 |||:::|||    :::|||
Db    1838 TGGACATCTTCACTCACTATGATATCCTCACCCCAAATGGCACCAAGGTG---------- 1887

Qy     584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604
                                                                  :::|||
Db    1888 ---------------------------------------------GCTGAGGGCCACA 1900

Qy     604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
                :::   |||
Db    1901 AAGCTAGTTTCTGT---------------------------------------------- 1914

Qy     624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644

Db    1914 ------------------------------------------------------------ 1914

Qy     644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
                    |||:::|||      |||         |||    |||         :::    |||     |
Db    1915 --------CTCGAAGACACTGAGTGTCAGGAGGATGTCTCCAAGCGGTATGAGTGTGCCA 1966

Qy     664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
             ||    |||                              :::    :::    |||    |||    |
Db    1967 ACTTTGGA----------------------GAGCAAGGCATCACTGTGGGTTGCTGGG 2002

Qy     684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
             ||  ::::::    |||    |||    :::    |||    :::||||||
Db    2003 ATCTCTACCGGCATGACATTGACTGTCAGTGGATTGACATCACGG 2047
```

RESULT 7
US-09-448-076-3
; Sequence 3, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: MNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-448-076-3

Alignment Scores:

Pred. No.:	6.12e-49	Length:	2262
Score:	634.50	Matches:	214
Percent Similarity:	40.03%	Conservative:	97
Best Local Similarity:	27.54%	Mismatches:	253
Query Match:	7.80%	Indels:	215
DB:	4	Gaps:	32

US-09-759-130B-381 (1-1453) x US-09-448-076-3 (1-2262)

```
Qy      30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr-------------- 44
           |||||||||        |||    |||    ||||||       |||:::
Db      40 CTGCTGCTG------TGCCTGCTGTGCAGTTCGTGCTTGGGGTCTCCGTCCCCTTCCACG 93

Qy      45 --------------------------AspLeuGluLeuArgLeuValAsn---GlyAsp 54
                                     |||         ||||||
Db      94 GGCCCTGAGAAGAAGGCCGGGAGCCAGGGGCTTCGGTTCCGGCTGGCTGGCTTCCCCAGG 153
```

```
Qy      55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db     154 AAGCCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGAT 213

Qy      75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
Db     214 GATGACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGC------TTCACA 267

Qy      95 PheAla---------------MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db     268 GAGGCCACAGGCTGGACCCACAGTGCCAAATATGGCCCTGGAACA---------GGCCGC 318

Qy     110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db     319 ATCTGGCTGGACAACTTGAGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCC 378

Qy     130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------ 147
Db     379 CGGGGCTGGGGGAACAGTGACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGAC 438

Qy     148 --------------TyrGlyGluAlaAsnLeu---------------------------- 153
Db     439 CAGCGCCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 498

Qy     154 --------------------------------GlyLeuArgLeuValAspGlyAsnAsn 162
Db     499 GAGGAGGTGCGAATTCGACCCGCCGTTGGGTGGGGCAGACGACCCCTGCCC--------- 549

Qy     163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
Db     550 GTGACGGAGGGGCTGGTGGAAGTCAGGCTTCCTGACGGCTGGTCGCAAGTGTGCGACAAA 609

Qy     183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
Db     610 GGCTGGAGCGCCCACAACAGCCACGTGGTCTGCGGGATGCTGGGCTTCCCCAGCGAAAAG 669

Qy     202 PheIleSerSerGlyValValAsn------SerProAlaValLeuArgProIleTrp-Le 219
Db     670 AGGGTCAACGC-GGCCTTCTACAGGCTGCTAGCCCAACGGCAGCAACACTCCTTTGGTCT 728

Qy     219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
Db     729 GCATGGGGTGGCGTGCGTGGGCACGGAGGCCCACCTCTCCCTCTGTTCCCTGGAGTTCTA 788

Qy     239 pGlyAsnHisAsp---------CysSerHisAsnGluAspValThrLeuThrCys----- 254
Db     789 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGAGCTGTGTGCC 848

Qy     255 ----------TyrAspSerSer-------------------------------------- 258
Db     849 AGGCCCTGTCTACGCGGCATCCAGTGGCCAGAAGAAGCAACAACAGTCGAAGCCTCAGGG 908

Qy     259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
Db     909 GGAGGCCCGTGTCCGTCTAAAGGGCGGCGCCCACCCTGGAGAGGGCCGGGTAGAAGTCCT 968

Qy     278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVa 298
Db     969 GAAGGCCAGCACATGGGGCACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1028

Qy     298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
Db    1029 GGTGTGTCGGGAGCTGGGCTTCGGGAGTGCTCGAGAAGCTCTGAGTGGCGCTCGCATGGG 1088

Qy     318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db    1089 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTCTGGACAGGAGCTCTCCCT 1148

Qy     338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
Db    1149 CTGGAAGTGCCCCCACAAGAACATCACAGCTGAGGATTGTTCACATAGCCAGGATGCCGG 1208

Qy     358 rValIleCysSer---------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
Db    1209 GGTCCGGTGCAACCTACCTTACACTGGGGCAGAGACCAGGATCCGACTCAGTGGGGGCCG 1268

Qy     375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----------GlnTrpTr 391
Db    1269 CAGCCAACATGAGGGGCGAGTCGAGGTGCAAATAGGGGGACCTGGGCCCCTTCGCTGGGG 1328

Qy     391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411
Db    1329 CCTCATCTGTGGGGATGACTGGGGGACCCTGGAGGCCATGGTGGCCTGTAGGCAACTGGG 1388

Qy     411 yCysProPheSerValPheGlySerArgArgAla------LysProSerAsnGluAlaAr 429
Db    1389 TCTGGGCTACGCCAACCACGGCCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1448

Qy     429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
Db    1449 AGAGGTGGTGATGAGTGGAGTGCGCTGCACAGGGACTGAGCTGTCCCTGGATCAGTGTGC 1508

Qy     449 rTyrAspGlyLysAlaLysArgThrCysPheArg---------ArgSerAspAlaGlyVa 466
Db    1509 CCATCATGGC---ACCCACATCACCTGCAAGAGGACAGGGACCCGCTTCACTGCTGGAGT 1565

Qy     466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486
Db    1566 CATCTGTTCTGAGACTGCATCA---GATCTGTTGCTG---------CACTCA-------- 1605

Qy     486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
Db    1606 ---------------------------------GCACTGGTGCA-GGAGACCGCCTACA 1630

Qy     506 rThrArgAsnAla---------AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
Db    1631 TCGAAGACCGGCCCCTGCATATGTTGTACTGTGC------------TGCGGAAGAGAACT 1678

Qy     523 t---------------HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
Db    1679 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGGTCACCGGCGTCTGCTCCGAT 1738

Qy     538 eTrpLeuAspAspVal-SerCys----------------------IleGlyAsnGluS 550
Db    1739 TCTCCTCCCAGATCCACAACCTGGGACGAGCTGACTTCAGGCCCAAGGCTGGGCGCCACT 1798

Qy     550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
Db    1799 CCTGGGTGTGGCACGAGTGC---------------------CATGGGCATTACCACAGCA 1837

Qy     569 luAspVal---------------IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db    1838 TGGACATCTTCACTCACTATGATATCCTCACCCCAAATGGCACCAAGGTG---------- 1887

Qy     584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604
Db    1888 ---------------------------------------------GCTGAGGGCCACA 1900

Qy     604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
Db    1901 AAGCTAGTTTCTGT---------------------------------------------- 1914

Qy     624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644
Db    1914 ------------------------------------------------------------ 1914

Qy     644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
Db    1915 --------CTCGAAGACACTGAGTGTCAGGAGGATGTCTCCAAGCGGTATGAGTGTGCCA 1966

Qy     664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db    1967 ACTTTGGA-----------------------GAGCAAGGCATCACTGTGGGTTGCTGGG 2002

Qy     684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
```

```
Db   2003 ATCTCTACCGGCATGACATTGACTGTCAGTGGATTGACATCACGG 2047

RESULT 8
US-09-702-572-3
; Sequence 3, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
;  APPLICANT:  Khodadoust, Mehran
;  TITLE OF INVENTION:  NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE:  MNI-073
;  CURRENT APPLICATION NUMBER: US/09/702,572
;  CURRENT FILING DATE:  2000-10-31
;  PRIOR APPLICATION NUMBER:  09/276,400
;  PRIOR FILING DATE:  1999-03-25
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-702-572-3

Alignment Scores:
Pred. No.:              6.12e-49      Length:        2262
Score:                  634.50        Matches:       214
Percent Similarity:     40.03%        Conservative:  97
Best Local Similarity:  27.54%        Mismatches:    253
Query Match:            7.80%         Indels:        215
DB:                     4             Gaps:          32

US-09-759-130B-381 (1-1453) x US-09-702-572-3 (1-2262)

Qy     30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr--------------- 44
          ||||||||       |||   |||  ||||||      |||:::
Db     40 CTGCTGCTG------TGCCTGCTGTGCAGTTCGTGCTTGGGGTCTCCGTCCCCTTCCACG 93

Qy     45 ---------------------------AspLeuGluLeuArgLeuValAsn---GlyAsp 54
                                        |||      ||||||
Db     94 GGCCCTGAGAAGAAGGCCGGGAGCCAGGGGCTTCGGTTCCGGCTGGCTGGCTTCCCCAGG 153

Qy     55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
             |||      |||   ||||||::::::         |||:::|||||||||:::||||||
Db    154 AAGCCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGAT 213

Qy     75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
          |||   :::        |||:::   ::::::|||::::::||||||         |||:::
Db    214 GATGACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGC------TTCACA 267

Qy     95 PheAla---------------MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
             |||                     ::::::|||                  |||:::
Db    268 GAGGCCACAGGCTGGACCCACAGTGCCAAATATGGCCCTGGAACA---------GGCCGC 318

Qy    110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
          ||||||||||||:::::::||||||      |||   |||    ::::::   ||||||
Db    319 ATCTGGCTGGACAACTTGAGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCC 378

Qy    130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------ 147
          |||    ||||||:::   :::|||   |||   ||||||    ||||||   |||
Db    379 CGGGGCTGGGGGAACAGTGACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGAC 438

Qy    148 ---------------TyrGlyGluAlaAsnLeu--------------------------- 153
                         :::   ::::::|||:::
Db    439 CAGCGCCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 498

Qy    154 ---------------------------------GlyLeuArgLeuValAspGlyAsnAsn 162
                                          |||   |||   :::
Db    499 GAGGAGGTGCGAATTCGACCCGCCGTTGGGTGGGGCAGACGACCCCTGCCC--------- 549

Qy    163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
                |||   |||||||||:::       :::   |||       :::||||||
Db    550 GTGACGGAGGGGCTGGTGGAAGTCAGGCTTCCTGACGGCTGGTCGCAAGTGTGCGACAAA 609

Qy    183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
          ||||||:::   :::   :::   |||||||||      ||||||   ||||||
Db    610 GGCTGGAGCGCCCACAACAGCCACGTGGTCTGCGGGATGCTGGGCTTCCCCAGCGAAAAG 669

Qy    202 PheIleSerSerGlyValValAsn------SerProAlaValLeuArgProIleTrp-Le 219
             ::::::   |||::::::          ||||||            :::|||  ||
Db    670 AGGGTCAACGC-GGCCTTCTACAGGCTGCTAGCCCAACGGCAGCAACACTCCTTTGGTCT 728

Qy    219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
           |     :::   |||   |||   |||         |||      |||          ::
Db    729 GCATGGGGTGGCGTGCGTGGGCACGGAGGCCCACCTCTCCCTCTGTTCCCTGGAGTTCTA 788

Qy    239 pGlyAsnHisAsp---------CysSerHisAsnGluAspValThrLeuThrCys----- 254
           :      :::|||          |||                  ::::::|||
Db    789 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGAGCTGTGTGCC 848

Qy    255 ----------TyrAspSerSer-------------------------------------- 258
                    |||   :::|||
Db    849 AGGCCCTGTCTACGCGGCATCCAGTGGCCAGAAGAAGCAACAACAGTCGAAGCCTCAGGG 908

Qy    259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
            :::      :::||||||   ||||||   :::        ||||||||||||:::
Db    909 GGAGGCCCGTGTCCGTCTAAAGGGCGGCGCCCACCCTGGAGAGGGCCGGGTAGAAGTCCT 968

Qy    278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVa 298
                         ||||||||||||      ||||||:::      ||||||   ||
Db    969 GAAGGCCAGCACATGGGGCACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1028

Qy    298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
           ||||||::::::||||||   |||:::|||                           :::
Db   1029 GGTGTGTCGGGAGCTGGGCTTCGGGAGTGCTCGAGAAGCTCTGAGTGGCGCTCGCATGGG 1088

Qy    318 nSerGlyGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
            |||         :::   |||      |||   |||||||||      |||   |||      ||
Db   1089 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTCTGGACAGGAGCTCTCCCT 1148

Qy    338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
           ||||   |||   |||                  ||||||   |||      |||
Db   1149 CTGGAAGTGCCCCCACAAGAACATCACAGCTGAGGATTGTTCACATAGCCAGGATGCCGG 1208

Qy    358 rValIleCysSer---------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
            |||   |||:::                ||||||:::      :::||||||:::   |||
Db   1209 GGTCCGGTGCAACCTACCTTACACTGGGGCAGAGACCAGGATCCGACTCAGTGGGGGCCG 1268

Qy    375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu------------GlnTrpTr 391
            :::            ||||||||||||||||:::|||              :::|||
Db   1269 CAGCCAACATGAGGGGCGAGTCGAGGTGCAAATAGGGGGACCTGGGCCCCTTCGCTGGGG 1328

Qy    391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411
             ||||||         :::|||      :::|||:::|||   |||:::||||||||
Db   1329 CCTCATCTGTGGGGATGACTGGGGGACCCTGGAGGCCATGGTGGCCTGTAGGCAACTGGG 1388

Qy    411 yCysProPheSerValPheGlySerArgArgAla------LysProSerAsnGluAlaAr 429
           |      ::::::      |||   :::                   |||
Db   1389 TCTGGGCTACGCCAACCACGGCCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1448

Qy    429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
           ::::::   ::::::   :::   |||||||||   |||   :::|||         |||
Db   1449 AGAGGTGGTGATGAGTGGAGTGCGCTGCACAGGGACTGAGCTGTCCCTGGATCAGTGTGC 1508

Qy    449 rTyrAspGlyLysAlaLysArgThrCysPheArg---------ArgSerAspAlaGlyVa 466
            :::   |||         ||||||   |||              |||      ||||||||
Db   1509 CCATCATGGC---ACCCACATCACCTGCAAGAGGACAGGGACCCGCTTCACTGCTGGAGT 1565

Qy    466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486
           ||||||||||:::   |||      ||||||   |||         ||||||
Db   1566 CATCTGTTCTGAGACTGCATCA---GATCTGTTGCTG---------CACTCA-------- 1605

Qy    486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
                                          |||      ||||||
Db   1606 ---------------------------------GCACTGGTGCA-GGAGACCGCCTACA 1630

Qy    506 rThrArgAsnAla---------AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
```

```
                    |||    |||              |||:::|||              ||||||:::   ::
Db     1631 TCGAAGACCGGCCCCTGCATATGTTGTACTGTGC------------TGCGGAAGAGAACT 1678

Qy      523 t--------------HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
            :              :::       :::      :::     ||||||   |||||
Db     1679 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGGTCACCGGCGTCTGCTCCGAT 1738

Qy      538 eTrpLeuAspAspVal-SerCys----------------------IleGlyAsnGluS 550
            |  |||  |||   :::                            |||           |
Db     1739 TCTCCTCCCAGATCCACAACCTGGGACGAGCTGACTTCAGGCCCAAGGCTGGGCGCCACT 1798

Qy      550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
            ||   :::|||   :::|||                    |||       |||
Db     1799 CCTGGGTGTGGCACGAGTGC--------------------CATGGGCATTACCACAGCA 1837

Qy      569 luAspVal---------------IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
              |||:::               |||:::|||   :::|||
Db     1838 TGGACATCTTCACTCACTATGATATCCTCACCCCAAATGGCACCAAGGTG---------- 1887

Qy      584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604
                                                            :::|||
Db     1888 ---------------------------------------------GCTGAGGGCCACA 1900

Qy      604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
              :::   |||
Db     1901 AAGCTAGTTTCTGT---------------------------------------------- 1914

Qy      624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644

Db     1914 ------------------------------------------------------------ 1914

Qy      644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
                    |||:::|||       |||      |||   |||       :::   |||   |
Db     1915 --------CTCGAAGACACTGAGTGTCAGGAGGATGTCTCCAAGCGGTATGAGTGTGCCA 1966

Qy      664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
            ||   |||                      :::   :::   |||   |||   |
Db     1967 ACTTTGGA-----------------------GAGCAAGGCATCACTGTGGGTTGCTGGG 2002

Qy      684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
            || ::::::   |||   |||   :::   |||   :::||||||
Db     2003 ATCTCTACCGGCATGACATTGACTGTCAGTGGATTGACATCACGG 2047
```

RESULT 9
US-09-276-400-1
; Sequence 1, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (143)..(2401)
US-09-276-400-1

Alignment Scores:

Pred. No.:	9.29e-49	Length:	2920
Score:	634.50	Matches:	214
Percent Similarity:	40.03%	Conservative:	97
Best Local Similarity:	27.54%	Mismatches:	253
Query Match:	7.80%	Indels:	215
DB:	3	Gaps:	32

US-09-759-130B-381 (1-1453) x US-09-276-400-1 (1-2920)

```
Qy       30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr--------------- 44
            |||||||||         |||   |||   ||||||      |||:::
Db      182 CTGCTGCTG------TGCCTGCTGTGCAGTTCGTGCTTGGGGTCTCCGTCCCCTTCCACG 235

Qy       45 --------------------------AspLeuGluLeuArgLeuValAsn---GlyAsp 54
                                         |||      ||||||
Db      236 GGCCCTGAGAAGAAGGCCGGGAGCCAGGGGCTTCGGTTCCGGCTGGCTGGCTTCCCCAGG 295

Qy       55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
               |||      |||      ||||||::::::        |||:::||||||||||:::||||||
Db      296 AAGCCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGAT 355

Qy       75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
            |||   :::          |||:::   ::::::|||::::::||||||      |||:::
Db      356 GATGACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGC------TTCACA 409

Qy       95 PheAla---------------MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
               |||                         ::::::|||                  |||:::
Db      410 GAGGCCACAGGCTGGACCCACAGTGCCAAATATGGCCCTGGAACA---------GGCCGC 460

Qy      110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
            ||||||||||||::::::||||||   |||   |||   ::::::   ||||||
Db      461 ATCTGGCTGGACAACTTGAGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCC 520

Qy      130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------ 147
            |||   ||||||:::   :::|||   |||   ||||||   ||||||   |||
Db      521 CGGGGCTGGGGGAACAGTGACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGAC 580

Qy      148 ---------------TyrGlyGluAlaAsnLeu--------------------------- 153
                           :::   ::::::|||:::
Db      581 CAGCGCCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 640

Qy      154 --------------------------------GlyLeuArgLeuValAspGlyAsnAsn 162
                                            |||   |||   :::
Db      641 GAGGAGGTGCGAATTCGACCCGCCGTTGGGTGGGGCAGACGACCCCTGCCC--------- 691

Qy      163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
                     |||   |||||||||:::         :::   |||      :::||||||
Db      692 GTGACGGAGGGGCTGGTGGAAGTCAGGCTTCCTGACGGCTGGTCGCAAGTGTGCGACAAA 751

Qy      183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
            ||||||:::   :::   :::   ||||||||      ||||||   ||||||
Db      752 GGCTGGAGCGCCCACAACAGCCACGTGGTCTGCGGGATGCTGGGCTTCCCCAGCGAAAAG 811

Qy      202 PheIleSerSerGlyValValAsn------SerProAlaValLeuArgProIleTrp-Le 219
               ::::::   |||::::::               ||||||            :::||| ||
Db      812 AGGGTCAACGC-GGCCTTCTACAGGCTGCTAGCCCAACGGCAGCAACACTCCTTTGGTCT 870

Qy      219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
            |      :::   |||   |||   |||         |||         |||         ::
Db      871 GCATGGGGTGGCGTGCGTGGGCACGGAGGCCCACCTCTCCCTCTGTTCCCTGGAGTTCTA 930

Qy      239 pGlyAsnHisAsp---------CysSerHisAsnGluAspValThrLeuThrCys----- 254
            :      :::|||         |||                        ::::::|||
Db      931 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGAGCTGTGTGCC 990

Qy      255 ---------TyrAspSerSer------------------------------------- 258
                     |||   :::|||
Db      991 AGGCCCTGTCTACGCGGCATCCAGTGGCCAGAAGAAGCAACAACAGTCGAAGCCTCAGGG 1050

Qy      259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
              :::      :::||||||      ||||||         :::         ||||||||||||:::
Db     1051 GGAGGCCCGTGTCCGTCTAAAGGGCGGCGCCCACCCTGGAGAGGGCCGGGTAGAAGTCCT 1110

Qy      278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVa 298
                           ||||||||||||||||         ||||||:::         ||||||   ||
Db     1111 GAAGGCCAGCACATGGGGCACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1170

Qy      298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
            ||||||::::::||||||   |||:::|||                              :::
Db     1171 GGTGTGTCGGGAGCTGGGCTTCGGGAGTGCTCGAGAAGCTCTGAGTGGCGCTCGCATGGG 1230
```

```
Qy     318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db    1231 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTCTGGACAGGAGCTCTCCCT 1290

Qy     338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
Db    1291 CTGGAAGTGCCCCCACAAGAACATCACAGCTGAGGATTGTTCACATAGCCAGGATGCCGG 1350

Qy     358 rValIleCysSer---------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
Db    1351 GGTCCGGTGCAACCTACCTTACACTGGGGCAGAGACCAGGATCCGACTCAGTGGGGGCCG 1410

Qy     375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu------------GlnTrpTr 391
Db    1411 CAGCCAACATGAGGGGCGAGTCGAGGTGCAAATAGGGGACCTGGGCCCCTTCGCTGGGG 1470

Qy     391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411
Db    1471 CCTCATCTGTGGGGATGACTGGGGGACCCTGGAGGCCATGGTGGCCTGTAGGCAACTGGG 1530

Qy     411 yCysProPheSerValPheGlySerArgArgAla------LysProSerAsnGluAlaAr 429
Db    1531 TCTGGGCTACGCCAACCACGGCCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1590

Qy     429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
Db    1591 AGAGGTGGTGATGAGTGGAGTGCGCTGCACAGGGACTGAGCTGTCCCTGGATCAGTGTGC 1650

Qy     449 rTyrAspGlyLysAlaLysArgThrCysPheArg---------ArgSerAspAlaGlyVa 466
Db    1651 CCATCATGGC---ACCCACATCACCTGCAAGAGGACAGGGACCCGCTTCACTGCTGGAGT 1707

Qy     466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486
Db    1708 CATCTGTTCTGAGACTGCATCA---GATCTGTTGCTG---------CACTCA-------- 1747

Qy     486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
Db    1748 --------------------------------GCACTGGTGCA-GGAGACCGCCTACA 1772

Qy     506 rThrArgAsnAla---------AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
Db    1773 TCGAAGACCGGCCCCTGCATATGTTGTACTGTGC-----------TGCGGAAGAGAACT 1820

Qy     523 t--------------HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
Db    1821 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGGTCACCGGCGTCTGCTCCGAT 1880

Qy     538 eTrpLeuAspAspVal-SerCys-----------------------IleGlyAsnGluS 550
Db    1881 TCTCCTCCCAGATCCACAACCTGGGACGAGCTGACTTCAGGCCCAAGGCTGGGCGCCACT 1940

Qy     550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
Db    1941 CCTGGGTGTGGCACGAGTGC--------------------CATGGGCATTACCACAGCA 1979

Qy     569 luAspVal---------------IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db    1980 TGGACATCTTCACTCACTATGATATCCTCACCCCAAATGGCACCAAGGTG---------- 2029

Qy     584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604
Db    2030 ---------------------------------------------GCTGAGGGCCACA 2042

Qy     604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
Db    2043 AAGCTAGTTTCTGT------------------------------------------- 2056

Qy     624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644
Db    2056 ------------------------------------------------------------ 2056

Qy     644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
Db    2057 --------CTCGAAGACACTGAGTGTCAGGAGGATGTCTCCAAGCGGTATGAGTGTGCCA 2108

Qy     664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db    2109 ACTTTGGA-----------------------GAGCAAGGCATCACTGTGGGTTGCTGGG 2144

Qy     684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db    2145 ATCTCTACCGGCATGACATTGACTGTCAGTGGATTGACATCACGG 2189
```

```
RESULT 10
US-09-448-076-1
; Sequence 1, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
;  APPLICANT: Khodadoust, Mehran et al.
;  TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
;  FILE REFERENCE: MNI-073CP
;  CURRENT APPLICATION NUMBER: US/09/448,076
;  CURRENT FILING DATE: 1999-11-23
;  EARLIER APPLICATION NUMBER: 60/117,580
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 09/276,400
;  EARLIER FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (143)..(2401)
US-09-448-076-1

Alignment Scores:
Pred. No.:              9.29e-49      Length:        2920
Score:                  634.50        Matches:       214
Percent Similarity:     40.03%        Conservative:  97
Best Local Similarity:  27.54%        Mismatches:    253
Query Match:            7.80%         Indels:        215
DB:                     4             Gaps:          32

US-09-759-130B-381 (1-1453) x US-09-448-076-1 (1-2920)

Qy      30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr-------------- 44
Db     182 CTGCTGCTG------TGCCTGCTGTGCAGTTCGTGCTTGGGGTCTCCGTCCCCTTCCACG 235

Qy      45 --------------------------AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db     236 GGCCCTGAGAAGAAGGCCGGGAGCCAGGGGCTTCGGTTCCGGCTGGCTGGCTTCCCCAGG 295

Qy      55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db     296 AAGCCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGAT 355

Qy      75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
Db     356 GATGACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGC------TTCACA 409

Qy      95 PheAla---------------MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db     410 GAGGCCACAGGCTGGACCCACAGTGCCAAATATGGCCCTGGAACA---------GGCCGC 460

Qy     110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db     461 ATCTGGCTGGACAACTTGAGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCC 520

Qy     130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------ 147
```

```
Db    521 CGGGGCTGGGGGAACAGTGACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGAC 580
Qy    148 --------------TyrGlyGluAlaAsnLeu-------------------------- 153
Db    581 CAGCGCCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 640
Qy    154 --------------------------------GlyLeuArgLeuValAspGlyAsnAsn 162
Db    641 GAGGAGGTGCGAATTCGACCCGCCGTTGGGTGGGGCAGACGACCCCTGCCC--------- 691
Qy    163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
Db    692 GTGACGGAGGGGCTGGTGGAAGTCAGGCTTCCTGACGGCTGGTCGCAAGTGTGCGACAAA 751
Qy    183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
Db    752 GGCTGGAGCGCCCACAACAGCCACGTGGTCTGCGGGATGCTGGGCTTCCCCAGCGAAAAG 811
Qy    202 PheIleSerSerGlyValValAsn------SerProAlaValLeuArgProIleTrp-Le 219
Db    812 AGGGTCAACGC-GGCCTTCTACAGGCTGCTAGCCCAACGGCAGCAACACTCCTTTGGTCT 870
Qy    219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
Db    871 GCATGGGGTGGCGTGCGTGGGCACGGAGGCCCACCTCTCCCTCTGTTCCCTGGAGTTCTA 930
Qy    239 pGlyAsnHisAsp---------CysSerHisAsnGluAspValThrLeuThrCys----- 254
Db    931 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGAGCTGTGTGCC 990
Qy    255 ----------TyrAspSerSer------------------------------------ 258
Db    991 AGGCCCTGTCTACGCGGCATCCAGTGGCCAGAAGAAGCAACAACAGTCGAAGCCTCAGGG 1050
Qy    259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
Db   1051 GGAGGCCCGTGTCCGTCTAAAGGGCGGCGCCCACCCTGGAGAGGGCCGGGTAGAAGTCCT 1110
Qy    278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVa 298
Db   1111 GAAGGCCAGCACATGGGGCACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1170
Qy    298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
Db   1171 GGTGTGTCGGGAGCTGGGCTTCGGGAGTGCTCGAGAAGCTCTGAGTGGCGCTCGCATGGG 1230
Qy    318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db   1231 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTCTGGACAGGAGCTCTCCCT 1290
Qy    338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
Db   1291 CTGGAAGTGCCCCCACAAGAACATCACAGCTGAGGATTGTTCACATAGCCAGGATGCCGG 1350
Qy    358 rValIleCysSer---------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
Db   1351 GGTCCGGTGCAACCTACCTTACACTGGGGCAGAGACCAGGATCCGACTCAGTGGGGGCCG 1410
Qy    375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----------GlnTrpTr 391
Db   1411 CAGCCAACATGAGGGGCGAGTCGAGGTGCAAATAGGGGACCTGGGCCCCTTCGCTGGGG 1470
Qy    391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411
Db   1471 CCTCATCTGTGGGGATGACTGGGGACCCTGGAGGCCATGGTGGCCTGTAGGCAACTGGG 1530
Qy    411 yCysProPheSerValPheGlySerArgArgAla------LysProSerAsnGluAlaAr 429
Db   1531 TCTGGGCTACGCCAACCACGGCCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1590
Qy    429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
Db   1591 AGAGGTGGTGATGAGTGGAGTGCGCTGCACAGGGACTGAGCTGTCCCTGGATCAGTGTGC 1650

Qy    449 rTyrAspGlyLysAlaLysArgThrCysPheArg---------ArgSerAspAlaGlyVa 466
Db   1651 CCATCATGGC---ACCCACATCACCTGCAAGAGGACAGGGACCCGCTTCACTGCTGGAGT 1707
Qy    466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486
Db   1708 CATCTGTTCTGAGACTGCATCA---GATCTGTTGCTG---------CACTCA-------- 1747
Qy    486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
Db   1748 ---------------------------------GCACTGGTGCA-GGAGACCGCCTACA 1772
Qy    506 rThrArgAsnAla---------AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
Db   1773 TCGAAGACCGGCCCCTGCATATGTTGTACTGTGC-----------TGCGGAAGAGAACT 1820
Qy    523 t--------------HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
Db   1821 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGGTCACCGGCGTCTGCTCCGAT 1880
Qy    538 eTrpLeuAspAspVal-SerCys----------------------IleGlyAsnGluS 550
Db   1881 TCTCCTCCCAGATCCACAACCTGGGACGAGCTGACTTCAGGCCCAAGGCTGGGCGCCACT 1940
Qy    550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
Db   1941 CCTGGGTGTGGCACGAGTGC--------------------CATGGGCATTACCACAGCA 1979
Qy    569 luAspVal--------------IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db   1980 TGGACATCTTCACTCACTATGATATCCTCACCCCAAATGGCACCAAGGTG---------- 2029
Qy    584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604
Db   2030 ---------------------------------------------GCTGAGGGCCACA 2042
Qy    604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
Db   2043 AAGCTAGTTTCTGT-------------------------------------------- 2056
Qy    624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644
Db   2056 ------------------------------------------------------------ 2056
Qy    644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
Db   2057 --------CTCGAAGACACTGAGTGTCAGGAGGATGTCTCCAAGCGGTATGAGTGTGCCA 2108
Qy    664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db   2109 ACTTTGGA-----------------------GAGCAAGGCATCACTGTGGGTTGCTGGG 2144
Qy    684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db   2145 ATCTCTACCGGCATGACATTGACTGTCAGTGGATTGACATCACGG 2189
```

RESULT 11
US-09-702-572-1
; Sequence 1, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/702,572
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```
;   LENGTH: 2920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (143)..(2401)
US-09-702-572-1

Alignment Scores:
Pred. No.:              9.29e-49      Length:         2920
Score:                  634.50        Matches:        214
Percent Similarity:     40.03%        Conservative:   97
Best Local Similarity:  27.54%        Mismatches:     253
Query Match:            7.80%         Indels:         215
DB:                     4             Gaps:           32

US-09-759-130B-381 (1-1453) x US-09-702-572-1 (1-2920)

Qy       30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr--------------- 44
             ||||||||        |||   |||   ||||||      |||::::
Db      182 CTGCTGCTG------TGCCTGCTGTGCAGTTCGTGCTTGGGGTCTCCGTCCCCTTCCACG 235

Qy       45 ---------------------------AspLeuGluLeuArgLeuValAsn---GlyAsp 54
                                           |||      ||||||
Db      236 GGCCCTGAGAAGAAGGCCGGGAGCCAGGGGCTTCGGTTCCGGCTGGCTGGCTTCCCCAGG 295

Qy       55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
               |||       |||   ||||||:::::    |||:::||||||||||||:::||||||
Db      296 AAGCCCTACGAGGGCCGCGTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGCGAT 355

Qy       75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
            |||    :::        |||:::   ::::::|||::::::||||||      |||:::
Db      356 GATGACTTCACGCTGCAGGCTGCCCACATCCTCTGCCGGGAGCTGGGC------TTCACA 409

Qy       95 PheAla---------------MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
               |||                      ::::::|||                |||:::
Db      410 GAGGCCACAGGCTGGACCCACAGTGCCAAATATGGCCCTGGAACA---------GGCCGC 460

Qy      110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
            |||||||||||::::::||||||    |||   |||    ::::::    ||||||
Db      461 ATCTGGCTGGACAACTTGAGCTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCCTCC 520

Qy      130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys------ 147
            |||     ||||||:::    :::|||   |||      ||||||     |||
Db      521 CGGGGCTGGGGGAACAGTGACTGTACGCACGATGAGGATGCTGGGGTCATCTGCAAAGAC 580

Qy      148 ---------------TyrGlyGluAlaAsnLeu--------------------------- 153
                           :::   ::::::|||:::
Db      581 CAGCGCCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 640

Qy      154 ---------------------------------GlyLeuArgLeuValAspGlyAsnAsn 162
                                             |||   |||   :::
Db      641 GAGGAGGTGCGAATTCGACCCGCCGTTGGGTGGGGCAGACGACCCCTGCCC--------- 691

Qy      163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
                     |||   ||||||||||:::        :::      |||      :::||||||
Db      692 GTGACGGAGGGGCTGGTGGAAGTCAGGCTTCCTGACGGCTGGTCGCAAGTGTGCGACAAA 751

Qy      183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
            ||||||:::   :::   :::   |||||||||         ||||||   ||||||
Db      752 GGCTGGAGCGCCCACAACAGCCACGTGGTCTGCGGGATGCTGGGCTTCCCCAGCGAAAAG 811

Qy      202 PheIleSerSerGlyValValAsn------SerProAlaValLeuArgProIleTrp-Le 219
               ::::::   |||::::::          ||||||            :::|||  ||
Db      812 AGGGTCAACGC-GGCCTTCTACAGGCTGCTAGCCCAACGGCAGCAACACTCCTTTGGTCT 870

Qy      219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
             |        :::   |||   |||   |||      |||      |||            ::
Db      871 GCATGGGGTGGCGTGCGTGGGCACGGAGGCCCACCTCTCCCTCTGTTCCCTGGAGTTCTA 930

Qy      239 pGlyAsnHisAsp---------CysSerHisAsnGluAspValThrLeuThrCys----- 254
             :      :::|||        |||               ::::::|||
Db      931 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGAGCTGTGTGCC 990

Qy      255 ---------TyrAspSerSer--------------------------------------- 258
                     |||   :::|||
Db      991 AGGCCCTGTCTACGCGGCATCCAGTGGCCAGAAGAAGCAACAACAGTCGAAGCCTCAGGG 1050

Qy      259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
              :::      :::||||||    ||||||    :::           ||||||||||||:::
Db     1051 GGAGGCCCGTGTCCGTCTAAAGGGCGGCGCCCACCCTGGAGAGGGCCGGGTAGAAGTCCT 1110

Qy      278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVa 298
                       |||||||||||||||        ||||||:::       ||||||    ||
Db     1111 GAAGGCCAGCACATGGGGCACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1170

Qy      298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
             ||||||:::::::||||||    |||:::|||                         :::
Db     1171 GGTGTGTCGGGAGCTGGGCTTCGGGAGTGCTCGAGAAGCTCTGAGTGGCGCTCGCATGGG 1230

Qy      318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
                |||      :::   |||        |||   |||||||||      |||       ||
Db     1231 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTCTGGACAGGAGCTCTCCCT 1290

Qy      338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
             ||||   |||   |||                  ||||||   |||      |||
Db     1291 CTGGAAGTGCCCCCACAAGAACATCACAGCTGAGGATTGTTCACATAGCCAGGATGCCGG 1350

Qy      358 rValIleCysSer---------AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
              |||   |||:::                ||||||:::      :::|||||||:::   |||
Db     1351 GGTCCGGTGCAACCTACCTTACACTGGGGCAGAGACCAGGATCCGACTCAGTGGGGGCCG 1410

Qy      375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----------GlnTrpTr 391
              :::            |||||||||||||||::::|||                  :::|||
Db     1411 CAGCCAACATGAGGGGCGAGTCGAGGTGCAAATAGGGGGACCTGGGCCCCTTCGCTGGGG 1470

Qy      391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411
                 ||||||       ::::|||        ::::|||:::|||   ||||:::|||||||||
Db     1471 CCTCATCTGTGGGGATGACTGGGGGACCCTGGAGGCCATGGTGGCCTGTAGGCAACTGGG 1530

Qy      411 yCysProPheSerValPheGlySerArgArgAla------LysProSerAsnGluAlaAr 429
             |      ::::::      |||   :::                   |||
Db     1531 TCTGGGCTACGCCAACCACGGCCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1590

Qy      429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
              ::::::   ::::::   :::   |||||||||   |||   :::|||      |||
Db     1591 AGAGGTGGTGATGAGTGGAGTGCGCTGCACAGGGACTGAGCTGTCCCTGGATCAGTGTGC 1650

Qy      449 rTyrAspGlyLysAlaLysArgThrCysPheArg---------ArgSerAspAlaGlyVa 466
              :::   |||          ||||||   |||                    |||||||||
Db     1651 CCATCATGGC---ACCCACATCACCTGCAAGAGGACAGGGACCCGCTTCACTGCTGGAGT 1707

Qy      466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486
             ||||||||||:::   |||     ||||||   |||            ||||||
Db     1708 CATCTGTTCTGAGACTGCATCA---GATCTGTTGCTG---------CACTCA-------- 1747

Qy      486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
                                                      |||      ||||||
Db     1748 ----------------------------------GCACTGGTGCA-GGAGACCGCCTACA 1772

Qy      506 rThrArgAsnAla---------AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
               |||   |||          |||:::|||             ||||||:::      ::
Db     1773 TCGAAGACCGGCCCCTGCATATGTTGTACTGTGC-----------TGCGGAAGAGAACT 1820

Qy      523 t---------------HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
            :                :::      :::      :::       ||||||    |||||
Db     1821 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGGTCACCGGCGTCTGCTCCGAT 1880

Qy      538 eTrpLeuAspAspVal-SerCys-----------------------IleGlyAsnGluS 550
             |   |||   |||   :::                             |||      |
Db     1881 TCTCCTCCCAGATCCACAACCTGGGACGAGCTGACTTCAGGCCCAAGGCTGGGCGCCACT 1940

Qy      550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
             ||   :::|||    :::|||                   |||         |||
Db     1941 CCTGGGTGTGGCACGAGTGC--------------------CATGGGCATTACCACAGCA 1979
```

```
Qy     569 luAspVal---------------IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
             |||:::               |||:::|||    :::|||
Db    1980 TGGACATCTTCACTCACTATGATATCCTCACCCCAAATGGCACCAAGGTG---------- 2029

Qy     584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604
                                                            :::|||
Db    2030 ----------------------------------------------GCTGAGGGCCACA 2042

Qy     604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
                 :::    |||
Db    2043 AAGCTAGTTTCTGT---------------------------------------------- 2056

Qy     624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644

Db    2056 ------------------------------------------------------------ 2056

Qy     644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
                   |||:::|||     |||      |||   |||       :::   |||   |
Db    2057 --------CTCGAAGACACTGAGTGTCAGGAGGATGTCTCCAAGCGGTATGAGTGTGCCA 2108

Qy     664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
             ||  |||                      :::   :::   |||   |||   |
Db    2109 ACTTTGGA----------------------GAGCAAGGCATCACTGTGGGTTGCTGGG 2144

Qy     684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
             || ::::::   |||    |||    :::    |||    :::||||||
Db    2145 ATCTCTACCGGCATGACATTGACTGTCAGTGGATTGACATCACGG 2189
```

RESULT 12
US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 28720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-341-587-7

Alignment Scores:

Pred. No.:	1.18e-45	Length:	28720
Score:	618.50	Matches:	386
Percent Similarity:	33.63%	Conservative:	183
Best Local Similarity:	22.81%	Mismatches:	501
Query Match:	7.60%	Indels:	629
DB:	4	Gaps:	80

US-09-759-130B-381 (1-1453) x US-09-341-587-7 (1-28720)

```
Qy      25 ValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGly--- 43
           :::|||      ::::::   ::::::          ||||||    |||
Db    2872 ATTGTTGACCTCCTGGTGGGGATGGATGAAGGGTTCTTGTGTTCCCCTGTAGGATCTGAA 2931

Qy      44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
           :::      |||   ||||||||||||||||||        |||   |||   |||||||||
Db    2932 TCCACTTTGGCCCTGAGACTGGTGAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGGTC 2991

Qy      64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThr 83
              :::||||||   ||||||||||||||||||||      |||:::|||         :::
Db    2992 CTATACCAAGGCTCCTGGGGCACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAC 3051

Qy      84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet---------PheArgPhe 100
           |||||||||:::||||||||||||     ::::::   :::          :::|||
Db    3052 GTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCAGCCCCAGGAAATGCCCAGTTT 3111

Qy     101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
           ||||||                 |||   |||   ||||||||||||    |||   |||:::
Db    3112 GGCCAGGGC---------TCAGGACCCATTGTCCTGGATGATGTGCGCTGCTCAGGACAC 3162

Qy     121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
           ||||||   ||||||   |||   |||         |||   |||||||||||||   |||
Db    3163 GAGTCTTACCTGTGGAGCTGCCCCCACAATGGCTGGCTCTCCCACAACTGTGGCCATCAT 3222

Qy     141 GluAspValGlyValAsnCysTyrGly------------------GluAlaAsnLeuGly 154
           ||||||   ||||||   |||   |||                          :::|||
Db    3223 GAAGATGCTGGTGTCATCTGCTCAGGTGGGCTTTCAAGACCTTGGGCTCCCTCTCTTAAG 3282

Qy     155 LeuArgLeuValAspGlyAsnAsn------------------------------------ 162
           |||:::              |||
Db    3283 TTGAAGTTTGCTCAGGAAGAAAATCCTAATTACATTCTGATCTCCTCACTCAAAGCTTTT 3342

Qy     163 -----------------------SerCysSerGlyArgValGluValLysPheGlnGlu 174
                                  ||||||               :::         ::::::
Db    3343 TCTATGTTTTCTATATTTCTGAAGTCTTGTTAGCTCTCTGCTAAGAATCTTTATGAATTT 3402

Qy     175 ArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArg 194
              :::   |||   |||                                       |||
Db    3403 TGCTACAGTACCTGGTGCAGC-------------------------------TGTGGC 3429

Qy     195 GlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla------ 212
              ||||||      |||            ::::::         ||||||:::
Db    3430 CACTTAGGCCAGGGCTCCGAACTGAA-ACAACAACCCAGACTTTATCCCCATCCTGAGGC 3488

Qy     213 --------------------------ValLeuArgProIleTrpLeuAspAspIleLeu 223
                                     ||||||   |||
Db    3489 AGTGCAAGGAAGAGGCAGAAGAGAAAAGTGCTGGCTCCC-------------------- 3527

Qy     224 CysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAsp 243
                    |||:::        :::|||   :::         |||:::         |||
Db    3528 ---CACGGCTCCATTTCTTCCCTGCTGAGTAGCACTGGTTGAGGGTATCGTGGACACAGC 3584

Qy     244 -----------------------CysSerHisAsnGluAspVal---ThrLeuThrCys 254
                                  ||||||   :::::::::   |||         |||
Db    3585 ACAGATAGCAGGGGCAGGGAGGGATCCTCTCACTACAAGGAACTGTGAACTAA-AGATGC 3643

Qy     255 TyrAspSerSer--------------------------------AspLeuGluLeuArg 263
                    |||                                 ::::::|||   :::
Db    3644 TTGTCTGGAAGTGGGTTCTCAGCTGAGACCCAGTGAGGAGGTCTGGAAATAGAGGCTCAA 3703

Qy     264 LeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrp 283
              :::|||         |||   |||               |||   :::   |||
Db    3704 GGGTTAGGAGT-GCAAATGCATGT--------------CTGTTTGTGTCAGGCCTGCTT 3747

Qy     284 GlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeu 303
           |||      :::                  |||||||||         :::|||   |||:::
Db    3748 GGAGTCCTTGACCTCAGGTCCTCTAAGAATGCTGCAGAGCACTGCCTGTGCCC-CAGGTC 3806

Qy     304 GlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySerAspVal 323
              ||||||                  |||   ||||||   :::||||||
Db    3807 TGGTGTGGG---------------GAGGGCAGCCCCCAT---GAGTCTGGCCAG------ 3842

Qy     324 ValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys------ 341
                                                    |||   |||
Db    3843 -----------------------------------------GCATGGCCTTGTTATTGC 3860

Qy     342 --------------------------ArgHis------------------------- 343
                                     ||||||
Db    3861 CTGTGGTCGGGGCTGTAAGATCACACAAGGCATTTGGGCTGGAGTGGCCTCCTCAGCCTT 3920

Qy     344 ------SerGlyThr--------------------------------ValAsnPheAsp 350
                 |||||||||                                ||||||
Db    3921 GCTGACTCAGGAACTGCCAAAACATCCAAGGGAGAGTGTTGGTTTTGGGTCAACCTGGTC 3980

Qy     351 CysLeuHisGlnAsnAspValSerValIleCysSerAsp------GlyAlaAspLeuGlu 368
           |||   ::::::   |||:::   :::   :::|||         |||   ||||||
Db    3981 ACCCTGGGCAGACACAAAGTTACTCACCTCGGAGCTGACAATAGTGGCCAGGATCTGCCT 4040
```

```
Qy    369 LeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGlu 388
                |||                              :::    |||  |||:::
Db   4041 GCACCCCTTATA---------------------TGGTGCATCTCTGTGGGAATTTACATG 4079

Qy    389 -GlnTrpTrpThrIleCysAspGln---------------AsnTrpLysAsnGluGlnAl 403
           |||      :::::|||||:::                 :::   :::      |||
Db   4080 GCAATGCCCCTCCCTCTGTGATAGGGACTAGGATGGACTGAGTGTCAGGCTTGCCCAGTT 4139

Qy    403 aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLy 423
           |||:::         |||      |||   |||
Db   4140 CCTTCTATCTTTGTTCCAGTTT---TGCCATTTTCTGTAT-------------------- 4176

Qy    423 sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe 443
                       |||   |||:::      :::|||:::|||
Db   4177 ---------AGTGCATCTGATCTGACCTTCTCTTCTCTTTCT------------------ 4209

Qy    443 rAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArgSerAs 463
                       |||:::            |||:::                |||
Db   4210 -------CACAGCTGCTCAGTCCCAGTCAACGCCCAG------------GCCAGGTGAGT 4250

Qy    463 pAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSe 483
                 :::::       |||         |||      :::|||         |||||
Db   4251 CCCCAGCATCCTTCATCGGGATGTCCCTTCTCTTTCTGCCCAGTTACCTCTTCCCCACTC 4310

Qy    483 rProCysTyrGlyArgLeuGluValLys-TyrGlnGly---------------------- 495
          |                     |||:::  |||   |||
Db   4311 CACAGAGCTCTCCTGCTTTTCTGTGCGGATACTGTGGGGCATATTATTTTTCCTCCCACC 4370

Qy    496 -----------GluTrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaV 512
                     |||            |||   :::|||         :::   |||:::|
Db   4371 ACTCTGTAACTGAGACCCCAGCATAGTGCTTCAACAGACATAGGGTTCAGGAGGCTTCTG 4430

Qy    512 alValCys-LysGlnLeu---GlyCysGlyLysProMetHisValPheGlyMetThrTyr 530
          ||   ||| :::   :::   |||      |||       |||            |||
Db   4431 TCTTCTGTTCAATTTATTTCAGGTTTCATGAAGCAGTTTCATACTGTACAATGGACAAGC 4490

Qy    531 PheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSer 550
                   :::         :::      ::::::::        |||     :::|||
Db   4491 CTTACAGGTTCAGGAAGTGGCCTCATGTTTAATGAGCTTTAGCTCATTTATATTCAGAGC 4550

Qy    551 AsnIle---------------------------------------TrpAspCysGluHis 557
             |||                                       |||:::
Db   4551 TG-ATACAACCTTTCTGAGAATTGAGAGTGACTGCCAAAGTCACCTGGGAAGGCAATGTC 4609

Qy    558 SerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleVal------------ 573
          |||         ||||||:::     |||   :::   :::   |||
Db   4610 AGTTCAAGCCTTAAACATGACTTCTGCCATGGGCAAGCAATGTCAGTGCAGGCCTGATAC 4669

Qy    573 ------------------------------------------------------------ 573

Db   4670 CTCCGTCCCTCACTCCTTCAAATACCCCAGATTTTGCAAGGCCATATCTGCATCCAGAAA 4729

Qy    574 ---------ThrCysSerGlyAspAlaThr---------------TrpGlyLeuArgLeu 585
                      |||||||||                        |||   |||   |||
Db   4730 AGGCAGAGGCCATGCTCGGGCAGGGAGAGGGATAATAAATATTTCTGGTGCCTCCACTTA 4789

Qy    586 ValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrp--- 604
                               ||||||   |||         :::         |||
Db   4790 ---------------------TCAGGAAACTTGATACCCCTTTGGGCGGCTCCTTGGTTC 4828

Qy    605 ------------GlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValVal 620
                         :::   |||               |||   :::      |||
Db   4829 CCCTAACATTTTAGCTCGAACTGT------------CAGAGTCTCAGCAATGGTGTCACA 4876

Qy    621 CysSerGlnLeuAsp-CysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerTh 640
          |||:::          |||                                       ||
Db   4877 TGTGCCCATCAGTCCATGT------------------------------------GTCAC 4900

Qy    640 rGlyTyrGlyLysIleTrpLeu------------AspAspValSerCysAspGlyAspGl 656
          ||||:::||||::::::  |||            :::|||:::|||
Db   4901 TGGATGGGGCAGGCTAACCCTTCGTAGCTGAGAAGAGGACATCTCACACTTCAGAGGTAG 4960

Qy    656 uSerAspLeuTrpSerCysArgAsnSer------GlyTrpGlyAsnAsnAspCysSerHi 674
              |||            :::               |||            ||||||
Db   4961 GAGGGATCGAACTGGTCTCCAGCAAGGCCTTTGTTCCTGG---------CTGTGCTCACT 5011

Qy    674 sSer-GluAspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValG 694
           ||| ||||||:::|||         |||         :::   |||:::
Db   5012 GAGCTGAAGACTTGGGTAGCACTTGGAGCAAGTGGCAGGAACCAGAAATTGAATAGTTT- 5070

Qy    694 lyGlySerSerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleL 714
                  |||  |||                     :::   |||:::
Db   5071 --------TCATGATGC----------------------TTGCCTGGTTCAGAGATTTTT 5101

Qy    714 euCysAla---------------------------------------------------- 716
            |||:::
Db   5102 TTTGTAGCTTTCCTCCCTCAAGTCTAATTTTGTCCTTTCTCTTTGTTGCAATTTACAGAT 5161

Qy    717 --------------AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuG 732
                                 |||:::::           ::::::          |||
Db   5162 ACTTGGCTGACCACCAACTTACCGGCATTGACAGTAGGTAAATAATCCTCTCGCCCCTCC 5221

Qy    732 luCysGlySerAlaIleArgValSerArg------------------------GluProH 744
                     :::::    :::   :::                        |||
Db   5222 CTAGGGCTCACTCTCTACCTCTGGACAAACGTTTCTTTTGAAAATGAAAGAATGAGGCTC 5281

Qy    744 isPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGly---------------- 758
                              |||:::      |||   |||
Db   5282 AAGCTGGCGCCTCTGTTTTTCATGTTTCCGCGAGTTGCCTGGGGAGGTAGACTCCCTGGG 5341

Qy    759 -----------------CysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpG 773
                                 ||||||          |||      |||::: |||:
Db   5342 AACCTAGTCCTGGGTCAGATGTTGGAGGCTGGAGGGTGCTGGTGACCTGTCTCCCCTGGG 5401

Qy    773 luTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisA 793
          ::                                     |||:::::||||||      |
Db   5402 AT------------------------------------TCTGTTTTATGTAGTCAGGAAA 5425

Qy    793 rgGlnProArgLeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisA 813
          ||                  :::          ||||||||:::
Db   5426 GA-----------------TCCTCAGCCAGGTGCTCAGGACAA----------------- 5451

Qy    813 laAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuC 833
                                              :::|||     :::   :::|||
Db   5452 -----------------------------------GCCCTGGAGGGCTCCCTAATCCTAC 5476

Qy    833 ysArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysG 853
                     :::                   |||         :::   :::||||||
Db   5477 TGGGACCTTGTTCCTGGCCCTCAGGCCACAGGCTGCAGACCTGCGAATAGTGGGGAAAG- 5535

Qy    853 lyAsnGlyLeuThrTrpAlaGluLysPheGlnCys------------------GluGlySerG 868
                                           |||                  |||
Db   5536 --------------------------------TGCCTGTCCCCATAGCTGTCTGGCCAA- 5562

Qy    868 luThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIleHisSerA 888
                           ||||||   :::             |||   :::   :::|
Db   5563 -----------------TGTCCTACCATT---------CCTGGCACTTTGCTAGAAGCCA 5596

Qy    888 rgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGlyLysSerG 908
          |||||      |||      ::::::       :::||||||
Db   5597 GAGAGGATCATGTGGGTGCCACCAAACTCCTGAACGTGAGG------------------- 5637

Qy    908 lnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrH 928
                                    ||||||      :::      ||||||
Db   5638 -----------------------CAGGTCTTGGCCTCCTATCTGGAGCTGTGC------- 5667

Qy    928 isTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuS 948
                           :::   :::|||||||
Db   5668 -----------------AGCTTCATCCTGTGTGAG------------------------- 5685

Qy    948 erThrThrGlyGlyLysTyrIleGlyGluArgSerValArg------------ValTrpG 964
                     |||   :::      |||       |||                  ||||
Db   5686 --------AATGGAGCTCACAATGAGGAGTGAGGCGTCCATTCCTGTCACCTCCAGTGGG 5737

Qy    964 lyHisArgPheHisCysLeuGlyAsnGlu---SerLeuLeuAsp---------------- 977
          ||||||||      |||      |||       :::   :::|||
```

```
Db   5738 GTCACAGG------TGCTTCCCCAATACTTGAGCTTCCATAGACTTGGGTGGAGTAGGAC 5791

Qy    978 --AsnCys-GlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
            :::|||        |||::::::             |||:::::::      :::
Db   5792 ATCACTGCTTCTTCCACTATCATGAAGCTGAACCTCTGTCTGTATTCATATTCAAAGGTG 5851

Qy    997 ValIleCysThr---------GlySerLeuThrGln-Pro-------------------- 1006
          :::   ||||||         |||   |||      |||
Db   5852 ATTACCTGCACACGTGACTTTGGCCAATTAGGAAGTGCCCTGAGTGTGGAACATTCCTTA 5911

Qy   1007 -LeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySe 1026
           :::        :::   ::::::   :::   :::|||   :::   |||   |||||
Db   5912 AATCCTTGACCTCATAATCAGTATGGATGAAGGGTTCTTGTGTTCC---CCTGTAGGATC 5968

Qy   1026 rAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAl 1046
          |                           ||||||||||||:::|||      ||||||
Db   5969 TGAA------------TCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGACAGGTGTCG 6016

Qy   1046 aGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAs 1066
           ||||||||||||:::   :::   |||   |||||||||:::|||||||||   |||||
Db   6017 AGGCCGAGTGGAGGTCCTGTATCGAGGCTCCTGGGGAACCGTGTGTGATGACAGCTGGGA 6076

Qy   1066 pLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaTh 1086
          |    :::|||||||:::|||||||||::::::|||||||||||    |||    :::|||
Db   6077 CACCAATGATGCCAATGTGGTCTGCAGGCAGCTGGGCTGTGGCTGGGCCATGTCGGCCCC 6136

Qy   1086 rValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysTh 1106
            :::|||   |||||:::|||||||||||||||   |||||||||:::     |||::
Db   6137 AGGAAATGCCCGGTTTGGCCAGGGCTCAGGACCCATTGTCCTGGATGATGTGCGCTGCTC 6196

Qy   1106 rGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysAr 1126
           :|||   ||||||:::||||||   ||||||   :::||||||      |||:::|||
Db   6197 AGGGAATGAGTCCTACCTGTGGAGCTGCCCCCACAAAGGCTGGCTCACCCACAACTGTGG 6256

Qy   1126 gHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGl 1146
            |||   |||||||||||||||||||||
Db   6257 CCATCACGAAGACGCTGGTGTCATCTGCTCA--------------------------- 6287

Qy   1146 uThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySe 1166
                                |||      :::   :::   :::   :::||||||
Db   6288 ------------------GGTGGGCC-TTCAAGACCTGGGGCTCCCTCTCTTGGGGTGG 6327

Qy   1166 rVal---GlyArgArgAsn--------------------IleThrThrAlaIleAlaGl 1178
           |||      |||||||||                     :::|||    :::    :::
Db   6328 AGTTTGCTCCAGAAGAAACTCCTAATTACATTCTGATCTCCTCACTCAAAGCTTCTCCTG 6387

Qy   1178 yIleValCys------------------------------------------------- 1181
            :::    |||
Db   6388 TGTTTCCTGTGTTTTTGAAGACTTGTTAGCTCTCTGCTAAGAATCCATATGAATTCACTG 6447

Qy   1182 ------------ArgGlnLeuGlyCysGlyGluAsnGlyValValSer----------- 1193
                              |||    |||:::            :::
Db   6448 CCTAGTGTTCCTGTGGTCACTTAGGACAGGGGACCAAACTGAAACAACAACCCAGACTTT 6507

Qy   1194 ---------------------------------------------LeuAlaPro----- 1196
                                                       |||||||||
Db   6508 ATCCCCTTCCTGAGGCAGTGCAAGGAAGAGGCAGAAGAAAAAATTTCTGGCTCCCCAGGG 6567

Qy   1197 ----------------LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGl 1211
                          ||||||    ||||||    |||:::   |||              ||
Db   6568 CTCCATTTCTCCCCTACTGAGTAGCACTGGGTGAGGGTATGG-TGGACACAGCACAGACA 6626

Qy   1211 nCysProLysThrHisIle---------------SerIleTrpGlnCysLeuSerAlaPr 1226
           |     :::    |||                     |||
Db   6627 GCGGGGCAGAGGAGGGATCCTCTCCTTAGGAGGAGGCTCATGGTAAGGAAAGGACATATG 6686

Qy   1226 oTrpGluArgArgIleSer------SerProAlaGluGluThrTrpIleThrCysGluAs 1244
           |||   |||    |||      :::|||          ||||||
Db   6687 TTGGGGTAGGGAATTCTCACTTGAGGCCCCAGTAAGGAGCATTTGGATT----------- 6735

Qy   1244 pArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAl 1264
                          ||||||                          ||||||
Db   6736 ----------------GGAGGC--------------------------ATATGGGCTAA 6752

Qy   1264 aGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCy 1284
            :::    |||    :::|||                                :::||
Db   6753 GAGTGCAAACGGGTATCTGTGC-------------------------------ATGTG 6779

Qy   1284 sGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGl 1304
          |                                                         ||
Db   6780 TGAC-----------------------------------------------CTGGG 6788

Qy   1304 yThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsn---------GluSerPh 1321
           |   |||          |||          |||                ::::::::
Db   6789 TCTTGGTCAGTTTCAGCTCCTTCCAAAGAACTGCAGTGCATTGTGTGTCTAGCAAGCATG 6848

Qy   1321 eLeuTrpAsp-----------------CysHisAlaLysPro---------------Tr 1330
           :   |||:::                  |||   |||||||||               ::
Db   6849 GTGTTGGGAGGGCAGCTCCCATGAGATCTGCCAGGCAAAGCCTTGTTATTACGTGTGGTT 6908

Qy   1330 pGlyGlnSerAspCysGlyHisLysGluAsp---------AlaGly------------- 1342
           :|||      |||      |||      |||         :::|||
Db   6909 TGGGATGGAAGAT-----CACACAGGGGATTTTTGCTGGAGTGGTATCCTGAGACTTGC 6962

Qy   1343 -------ValArgCys---------------SerGlyGlnSerLeuLysSerLeuAsnAl 1355
                 |||   |||                     |||      |||   |||:::
Db   6963 TGACTTGGTGAATTGCTAAAACCTGCAAGGGAAAGGGTTGGTTTTGGTTCAATTGGACAT 7022

Qy   1355 aSerSerGlyHis-----------------------------------------LeuAl 1361
            :::|||   |||                                         |||
Db   7023 CACATCCAGACACAGAGTTAATTGCCTTGGAGCTGGAAATGGTGGACAGAATCTGCTTGG 7082

Qy   1361 aLeuIleLeuSerSerIlePhe 1368
             ::::::   |||:::|||
Db   7083 ATCCCTTATAAGGAGCGTCTTT 7104
```

RESULT 13
US-08-477-674-9
; Sequence 9, Application US/08477674
; Patent No. 5644035
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert F.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Casipit, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,674
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE:
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470

```
;       TELEX:  N/A
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2285 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-477-674-9

Alignment Scores:
Pred. No.:               3.56e-19       Length:        2285
Score:                   314.00         Matches:       107
Percent Similarity:      42.00%         Conservative:  40
Best Local Similarity:   30.57%         Mismatches:    108
Query Match:             3.86%          Indels:        96
DB:                      1              Gaps:          15

US-09-759-130B-381 (1-1453) x US-08-477-674-9 (1-2285)

Qy    1031 LeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGlu 1050
             :::   |||      :::||||||  ||||||   :::       ||||||||||||
Db     234 GTGAATGATGGTGACATGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAG 293

Qy    1051 IleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAla 1070
           |||::::::  |||   ||||||||::::|||||||:::  ||||||||||::::||||||
Db     294 ATCTTCTACAGAGGCCAGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCC 353

Qy    1071 HisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHis 1090
              |||||||||:::   ||||||           |||       |||        |||
Db     354 AGCGTCGTCTGCCGGGCCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCC 413

Qy    1091 PheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSer 1110
           ||||||:::||||||||||||||||||  ||||||::::::  ||||||||||||||||:::
Db     414 TTCGGGCAAGGATCAGGCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCC 473

Qy    1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130
              |||      |||   |||   ||||||   :::    :::||||||||||:::    |||
Db     474 TCACTGGCCGACTGCAAGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGAC 533

Qy    1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150
           |||||||||:::|||:::        |||                   |||   |||
Db     534 GCTGGTGTGGTCTGCACCAATGAAACC---------------AGGAGCACCCACACCCTG 578

Qy    1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170
                 :::                                                |||
Db     579 GACCTCTCC---------------------------------------------------AGG 590

Qy    1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190
              ::::::   |||:::           |||        |||   ||||||        :::
Db     591 GAGCTCTCGGAGGCCCTTGGCCAGATCTTTGACAGCCAGCGGGGCTGCGACCTGTCC--- 647

Qy    1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIle 1210
              :::|||:::          :::         ||||||
Db     648 ---ATCAGCGTGAATGTGCAGGGCGAGGACGCCCTGGGCTTC------------------ 686

Qy    1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230
              |||      |||   |||                |||:::|||       |||
Db     687 ---TGTGGCCACACGGTCATC---------------CTGACTGCCAACCTGGAG------ 722

Qy    1231 IleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250
                       |||:::      |||                  :::   :::
Db     723 -----------GCCCAGGCCCTGTGGAAGGAGCCGGGCAGCAATGTCACCATG---AGT 767

Qy    1251 GlyAspThrGluCys-----------------------------------SerGlyArg 1258
             |||   ||||||                                   |||   |||
Db     768 GTGGATGCTGAGTGTGTGCCCATGGTCAGGGACCTTCTCAGGTACTTCTACTCCCGAAGG 827

Qy    1259 ValGluIle--------------------TrpHisAlaGlySerTrpGlyThrValCys 1271
           :::::::|||                   :::|||
Db     828 ATTGACATCACCCTGTCGTCAGTCAAGTGCTTCCAC----------------------- 863

Qy    1272 AspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly---CysGly 1290
                                    ||||||   |||               |||    |||   |||
Db     864 ------------AAGCTGGCCTCTGCCTATGGGGCCAGGCAGCTGCAGGGCTACTGCGCA 911

Qy    1291 SerAlaLeuAlaAlaLeu----ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308
             |||       |||   |||          :::|||   ||||||               |||   :::
Db     912 AGCCTCTTTGCCATCCTCCTCCCCCAGGACCCCTCGTTCCA-GATGCCCCTGGACCTGTA 970

Qy    1309 TrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328
              |||                |||   |||:::                        :::
Db     971 TGCCTA-----------TGCAGTGGCCAC--------------------AGGGGACGC 997

Qy    1329 ProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348
             |||   |||                           :::|||               |||
Db     998 CCTGCTGGA----------------------GAAGCTCTGCCTACAGTTCCTGGCCTG 1033

Qy    1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358
              |||:::   |||:::|||     |||
Db    1034 GAACTTCGAGGCCTTGACGCAGGCCGAGGC 1063
```

```
RESULT 14
US-08-473-791-9
; Sequence 9, Application US/08473791
; Patent No. 5736340
;  GENERAL INFORMATION:
;    APPLICANT:  Koths, Kirston E.
;    APPLICANT:  Halenbeck, Robert F.
;    APPLICANT:  Taylor, Eric W.
;    APPLICANT:  Wang, Alice M.
;    APPLICANT:  Casipit, Clayton L.
;    TITLE OF INVENTION:  Secreted Mac-2-Binding Glycoprotein
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cetus Oncology Corporation
;      STREET:  1400 Fifty-Third Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,791
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/316,714
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/961,404
;      FILING DATE:  15-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goldman, Kenneth M.
;      REGISTRATION NUMBER:  34,174
;      REFERENCE/DOCKET NUMBER:  2595.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 420-3152
;      TELEFAX:  (510) 658-5470
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-473-791-9

Alignment Scores:
Pred. No.:               3.56e-19       Length:        2285
Score:                   314.00         Matches:       107
Percent Similarity:      42.00%         Conservative:  40
```

```
Best Local Similarity:  30.57%        Mismatches:    108
Query Match:            3.86%         Indels:        96
DB:                     1             Gaps:          15

US-09-759-130B-381 (1-1453) x US-08-473-791-9 (1-2285)

Qy   1031 LeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGlu 1050
          :::  |||        :::||||||  ||||||  :::           |||||||||||
Db    234 GTGAATGATGGTGACATGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAG 293

Qy   1051 IleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAla 1070
          |||:::::  |||  ||||||||:::||||||:::   ||||||||:::||||||
Db    294 ATCTTCTACAGAGGCCAGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCC 353

Qy   1071 HisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHis 1090
             |||||||||:::  ||||||        |||      |||        |||
Db    354 AGCGTCGTCTGCCGGGCCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCC 413

Qy   1091 PheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSer 1110
          ||||||:::|||||||||||||||  ||||||::::::  ||||||||||||||||:::
Db    414 TTCGGGCAAGGATCAGGCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCC 473

Qy   1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130
             |||      |||  |||  ||||||  :::   :::||||||||||:::   |||
Db    474 TCACTGGCCGACTGCAAGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGAC 533

Qy   1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150
          |||||||||:::|||      |||                     |||   |||
Db    534 GCTGGTGTGGTCTGCACCAATGAAACC---------------AGGAGCACCCACACCCTG 578

Qy   1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170
             :::                                                  |||
Db    579 GACCTCTCC---------------------------------------------AGG 590

Qy   1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190
             ::::::  |||:::      |||          |||   ||||||      :::
Db    591 GAGCTCTCGGAGGCCCTTGGCCAGATCTTTGACAGCCAGCGGGGCTGCGACCTGTCC--- 647

Qy   1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIle 1210
             :::|||:::            :::        ||||||
Db    648 ---ATCAGCGTGAATGTGCAGGGCGAGGACGCCCTGGGCTTC------------------ 686

Qy   1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230
             |||      |||   |||                |||:::|||  |||
Db    687 ---TGTGGCCACACGGTCATC--------------CTGACTGCCAACCTGGAG------ 722

Qy   1231 IleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250
                   |||:::    |||                :::    :::
Db    723 -----------GCCCAGGCCCTGTGGAAGGAGCCGGGCAGCAATGTCACCATG---AGT 767

Qy   1251 GlyAspThrGluCys-----------------------------------SerGlyArg 1258
             |||  ||||||                                      |||  |||
Db    768 GTGGATGCTGAGTGTGTGCCCATGGTCAGGGACCTTCTCAGGTACTTCTACTCCCGAAGG 827

Qy   1259 ValGluIle--------------------TrpHisAlaGlySerTrpGlyThrValCys 1271
          ::::::|||                     :::|||
Db    828 ATTGACATCACCCTGTCGTCAGTCAAGTGCTTCCAC------------------------ 863

Qy   1272 AspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly---CysGly 1290
                      ||||||   |||            |||    |||    |||
Db    864 -----------AAGCTGGCCTCTGCCTATGGGGCCAGGCAGCTGCAGGGCTACTGCGCA 911

Qy   1291 SerAlaLeuAlaAlaLeu------ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308
          |||       |||  |||      :::|||  ||||||            |||    :::
Db    912 AGCCTCTTTGCCATCCTCCTCCCCAGGACCCCTCGTTCCA-GATGCCCCTGGACCTGTA 970

Qy   1309 TrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328
             |||            |||  |||:::                           :::
Db    971 TGCCTA------------TGCAGTGGCCAC--------------------AGGGGACGC 997

Qy   1329 ProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348
          |||  |||                       :::|||                  |||
Db    998 CCTGCTGGA-----------------------GAAGCTCTGCCTACAGTTCCTGGCCTG 1033

Qy   1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358
                |||:::  |||:::|||     |||
Db   1034 GAACTTCGAGGCCTTGACGCAGGCCGAGGC 1063
```

```
RESULT 15
US-08-316-714-9
; Sequence 9, Application US/08316714
; Patent No. 5965382
;  GENERAL INFORMATION:
;    APPLICANT:  Koths, Kirston E.
;    APPLICANT:  Halenbeck, Robert F.
;    APPLICANT:  Taylor, Eric W.
;    APPLICANT:  Wang, Alice M.
;    APPLICANT:  Casipit, Clayton L.
;    TITLE OF INVENTION:  Secreted Mac-2-Binding Glycoprotein
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cetus Oncology Corporation
;      STREET:  1400 Fifty-Third Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/316,714
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/961,404
;      FILING DATE:  15-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goldman, Kenneth M.
;      REGISTRATION NUMBER:  34,174
;      REFERENCE/DOCKET NUMBER:  2595.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 420-3152
;      TELEFAX:  (510) 658-5470
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-316-714-9

Alignment Scores:
Pred. No.:              3.56e-19      Length:         2285
Score:                  314.00        Matches:        107
Percent Similarity:     42.00%        Conservative:   40
Best Local Similarity:  30.57%        Mismatches:     108
Query Match:            3.86%         Indels:         96
DB:                     2             Gaps:           15

US-09-759-130B-381 (1-1453) x US-08-316-714-9 (1-2285)

Qy   1031 LeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGlu 1050
          :::  |||        :::||||||  ||||||  :::           |||||||||||
Db    234 GTGAATGATGGTGACATGCGGCTGGCCGATGGGGGCGCCACCAACCAGGGCCGCGTGGAG 293

Qy   1051 IleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAla 1070
          |||:::::  |||  ||||||||:::||||||:::   ||||||||:::||||||
Db    294 ATCTTCTACAGAGGCCAGTGGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCC 353

Qy   1071 HisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHis 1090
             |||||||||:::  ||||||        |||      |||        |||
```

```
Db     354 AGCGTCGTCTGCCGGGCCCTGGGCTTCGAGAACGCCACCCAGGCTCTGGGCAGAGCTGCC 413

Qy    1091 PheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSer 1110
           ||||||:::||||||||||||||    ||||||::::::   ||||||||||||||:::
Db     414 TTCGGGCAAGGATCAGGCCCCATCATGCTGGACGAGGTCCAGTGCACGGGAACCGAGGCC 473

Qy    1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130
              |||      |||   |||   ||||||   :::   :::|||||||||:::   |||
Db     474 TCACTGGCCGACTGCAAGTCCCTGGGCTGGCTGAAGAGCAACTGCAGGCACGAGAGAGAC 533

Qy    1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150
           |||||||||:::|||:::        |||                  |||   |||
Db     534 GCTGGTGTGGTCTGCACCAATGAAACC---------------AGGAGCACCCACACCCTG 578

Qy    1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170
                 :::                                                |||
Db     579 GACCTCTCC---------------------------------------------AGG 590

Qy    1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190
              ::::::   |||:::      |||         |||   ||||||      :::
Db     591 GAGCTCTCGGAGGCCCTTGGCCAGATCTTTGACAGCCAGCGGGGCTGCGACCTGTCC--- 647

Qy    1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIle 1210
              :::|||:::         :::         ||||||
Db     648 ---ATCAGCGTGAATGTGCAGGGCGAGGACGCCCTGGGCTTC------------------ 686

Qy    1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230
              |||      |||   |||         |||:::|||      |||
Db     687 ---TGTGGCCACACGGTCATC---------------CTGACTGCCAACCTGGAG------ 722

Qy    1231 IleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250
                       |||:::      |||                  :::   :::
Db     723 ------------GCCCAGGCCCTGTGGAAGGAGCCGGGCAGCAATGTCACCATG---AGT 767

Qy    1251 GlyAspThrGluCys------------------------------------SerGlyArg 1258
              |||   ||||||                                    |||   |||
Db     768 GTGGATGCTGAGTGTGTGCCCATGGTCAGGGACCTTCTCAGGTACTTCTACTCCCGAAGG 827

Qy    1259 ValGluIle---------------------TrpHisAlaGlySerTrpGlyThrValCys 1271
           ::::::|||                     :::|||
Db     828 ATTGACATCACCCTGTCGTCAGTCAAGTGCTTCCAC------------------------ 863

Qy    1272 AspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly---CysGly 1290
                       ||||||   |||               ||||      |||   |||
Db     864 ------------AAGCTGGCCTCTGCCTATGGGGCCAGGCAGCTGCAGGGCTACTGCGCA 911

Qy    1291 SerAlaLeuAlaAlaLeu------ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308
           |||      |||   |||      :::|||   ||||||            |||   :::
Db     912 AGCCTCTTTGCCATCCTCCTCCCCAGGACCCCTCGTTCCA-GATGCCCCTGGACCTGTA 970

Qy    1309 TrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328
              |||            |||   |||:::                           :::
Db     971 TGCCTA------------TGCAGTGGCCAC--------------------AGGGGACGC 997

Qy    1329 ProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348
           |||   |||                        :::|||               |||
Db     998 CCTGCTGGA-----------------------GAAGCTCTGCCTACAGTTCCTGGCCTG 1033

Qy    1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358
              |||:::   |||:::|||   |||
Db    1034 GAACTTCGAGGCCTTGACGCAGGCCGAGGC 1063
```

Search completed: May 12, 2003, 19:17:06
Job time : 236 secs